

KINH NGHIỆM THIẾT KẾ MẪU CHỦ

cho các cuộc điều tra hộ gia đình của Bangladesh

(Tiếp theo)

3. Kinh nghiệm lựa chọn mẫu để có thể kiểm soát hiệu quả thiết kế của các cuộc điều tra phức tạp

Bảng 6: Đặc trưng thống kê của các đơn vị phân lớp theo thành thị/nông thôn

Vùng	Đơn vị phân tầng	Thành thị/ Nông thôn	Nhỏ nhất	Trung vị	Trung bình	Lớn nhất	Độ lệch chuẩn
Barisal	PStrong ⁴	Nông thôn	0	0.99	2.93	100	7.11
		Thành thị	0	14.93	25.37	100	26.54
	PAgri ⁵	Nông thôn	0	61.68	59.75	100	23.59
		Thành thị	0	7.75	20.33	100	24.91
	POal ⁶	Nông thôn	0	69.46	66.26	100	22.92
		Thành thị	0	50.57	50.46	100	23.62
Chittagong	PStrong	Nông thôn	0	4.05	7.46	100	10.70
		Thành thị	0	30.48	38.11	100	31.68
	PAgri	Nông thôn	0	46.94	48.99	100	25.85
		Thành thị	0	4.55	15.21	100	22.29
	POal	Nông thôn	0	58.33	57.26	100	22.87
		Thành thị	0	38.63	40.41	100	25.21
Dhaka	PStrong	Nông thôn	0	1.85	5.37	100	9.49
		Thành thị	0	57.56	53.90	100	35.59
	PAgri	Nông thôn	0	67.42	62.93	100	24.50
		Thành thị	0	1.25	10.24	100	19.36
	POal	Nông thôn	0	61.54	61.37	100	20.77
		Thành thị	0	48.54	48.32	100	26.22
Khulna	PStrong	Nông thôn	0	15.27	17.87	100	14.37
		Thành thị	0	44.17	46.28	100	27.19
	PAgri	Nông thôn	0	71.07	65.90	100	22.95
		Thành thị	0	6.49	19.61	100	25.59
	POal	Nông thôn	0	61.17	60.87	100	20.36
		Thành thị	0	43.33	44.54	100	22.60

⁴ Tỷ lệ hộ gia đình có nhà ở vật liệu tốt

⁵ Tỷ lệ hộ gia đình có sản xuất nông nghiệp là nguồn thu nhập chính

⁶ Tỷ lệ hộ gia đình sở hữu đất nông nghiệp

Vùng	Đơn vị phân tầng	Thành thị/ Nông thôn	Nhỏ nhất	Trung vị	Trung bình	Lớn nhất	Độ lệch chuẩn
Rajshahi	PStrong	Nông thôn	0	3.80	7.68	100	11.05
		Thành thị	0	33.33	39.39	100	30.84
	PAgri	Nông thôn	0	76.09	70.46	100	22.39
		Thành thị	0	12.00	24.86	100	27.29
	POal	Nông thôn	0	57.03	57.14	100	19.51
		Thành thị	0	39.39	40.72	100	20.87
Sylhet	PStrong	Nông thôn	0	11.83	17.92	100	18.68
		Thành thị	0	49.07	47.42	100	29.50
	PAgri	Nông thôn	0	58.76	56.09	100	28.57
		Thành thị	0	7.25	18.55	100	23.43
	POal	Nông thôn	0	49.38	49.36	100	22.63
		Thành thị	0	38.65	40.68	100	22.71
Bangladesh	PStrong	Tất cả	0	6.06	17.47	100	25.36
	PAgri	Tất cả	0	56.82	51.12	100	31.86
	POal	Tất cả	0	56.43	55.64	100	23.06

Nguồn: Tính toán của tác giả sử dụng dữ liệu Tổng điều tra dân số năm 2001 tiến hành bởi BBS.

Một vài phát hiện chỉ ra rằng các phân loại nông thôn/thành thị cần được xem xét một cách cẩn thận. Chi tiết, có các PSU ở khu vực đô thị, trong đó tất cả các hộ gia đình có sản xuất nông nghiệp là nguồn thu nhập chính trong khi có những PSU ở các khu vực nông thôn có một số hộ gia đình nguồn thu nhập chính không từ sản xuất nông nghiệp. Bảng 6 cũng cho thấy rằng quyền sở hữu đất nông nghiệp không phải là một yếu tố phân biệt tốt cho khu vực thành thị /nông thôn. Điều này có thể cho thấy rằng có rất nhiều chủ sở hữu tại các khu vực thành thị, thuê hay cho thuê đất nông nghiệp của họ và do đó, làm giảm giá trị của đơn vị phân tầng Poal.

Xem xét các đặc trưng thống kê: độ lệch chuẩn, giá trị nhỏ nhất, giá trị trung bình và giá trị lớn nhất, thấy rằng tỷ lệ hộ có nhà ở vật liệu tốt (Pstrong) không thay đổi nhiều trong khu vực nông thôn. Tính trung bình, các hộ gia đình có nhà ở vật liệu tốt có tỷ lệ thấp hơn đáng kể ở các vùng nông thôn. Mặt khác, mặc dù các đặc trưng thống kê (giá trị nhỏ nhất,

trung bình, lớn nhất) của PAgri là tương đối giống nhau cho các khu vực thành thị và nông thôn ở một số vùng, số lượng các hộ gia đình với nông nghiệp là nguồn thu nhập chính là thấp hơn đáng kể hơn trong các khu vực thành thị, trên mức trung bình. Những kết quả này nhắc nhở chúng ta phân lớp các khu vực thành thị sử dụng đơn vị phân lớp là PStrong và khu vực nông thôn sử dụng đơn vị phân lớp là PAgri. Đặc biệt, số lượng hộ gia đình và PSU ở khu vực nông thôn là nhiều hơn gấp hai lần khu vực thành thị, và tương ứng với 4 lớp và 2 lớp đã được tính trước cho khu vực nông thôn và khu vực thành thị. Ranh giới lớp đầu tiên được thiết lập như là các tứ phân vị của PAgri cho các khu vực nông thôn và trung bình của PStrong cho các khu vực thành thị. Tuy nhiên, các lớp nhỏ hoặc lớp có tổng số hộ gia đình ít hơn khoảng lấy mẫu của lớp được kết hợp với các lớp liền kề.

Nói chung, những lợi thế quan trọng quy trình phân lớp (ngầm) ở đây là đơn giản để thực hiện và cho ra kết quả có thể chấp nhận

được. Tuy nhiên, các nghiên cứu trong tương lai có thể xem xét việc thực hiện các thủ tục phân lớp tối ưu hơn như những biện pháp mà Sethiv (1963) và Kozak (2004) đề xuất.

3.2. Chọn mẫu

Để hiệu quả thiết kế được kiểm soát thì phải đảm bảo rằng các quyền số điều tra trong các tầng không thay đổi nhiều. Một sự khác biệt lớn về quyền số trong một tầng sẽ không làm gia tăng phương sai của các ước lượng, do đó, các nhà điều tra thống kê thường lựa chọn giữ nguyên quyền số cơ sở gần như tương tự trong một tầng. Do quyền số cơ sở là nghịch đảo của xác suất chọn một đơn vị lấy mẫu cuối cùng, sau đó duy trì quyền số cơ sở tương tự hoặc gần như thống nhất là tương đương với việc duy trì xác suất lựa chọn giống nhau hoặc gần như tương tự nhau trong một tầng. Phần này thảo luận về cách làm thế nào để đạt được mục đích. Ở đây, chúng ta đề xuất một thiết kế chọn mẫu đơn giản hai giai đoạn trong một tầng d : (i) PSU _{α} sẽ được lựa chọn với xác suất tỷ lệ với quy mô và (ii) hộ gia đình β từ PSU _{α} sẽ được lựa chọn bằng cách lấy mẫu ngẫu nhiên đơn giản hoặc chọn mẫu hệ thống, trong tầng d mà tất cả các PSU cũng được nhóm lại bằng đơn vị phân tầng tầng ngầm. Vì vậy, trong phạm vi của tầng d và tầng (ngầm) h , xác suất f_d mà một hộ gia đình được chọn từ một PSU sẽ là:

$$f_d = \frac{n'_d}{M_d} \quad (3)$$

Trong đó n'_d là tổng số cỡ mẫu cho tầng d được định nghĩa ở cột cuối cùng của bảng 5 (Kish Allocation, Index = 1), M_d là đơn vị xác định quy mô của tầng d (ví dụ, tổng số hộ gia đình bình quân vùng xác định trong Tổng điều tra dân số năm 2001) và $M_{h\alpha}$ là đơn vị đo quy mô của PSU _{α} tại tầng h (tức là, tổng số hộ gia

đình cho PSU _{α} của tầng h).

Trong một thiết kế chọn mẫu chùm hai giai đoạn:

$$f_d = P(\alpha\beta|h) = P(\alpha|h)P(\beta|h\alpha) \quad (4)$$

trong đó $P(\alpha|h)$ là xác suất lựa chọn của PSU _{α} và $P(\beta|h\alpha)$ là xác suất lựa chọn hộ gia đình β thuộc PSU _{α} tại tầng h được chọn. Do đó:

$$f_d = \frac{a_h M_{h\alpha}}{\sum M_{h\alpha}} \frac{b_h}{M_{h\alpha}} = \frac{a_h b_h}{\sum M_{h\alpha}} \quad (5)$$

Trong đó a_h là số lượng PSU được lấy mẫu từ tầng h , và b_h là số hộ gia đình được lựa chọn từ tầng h .

Thuật ngữ $P(h\beta/a) = \frac{b_h}{M_{h\alpha}}$ đại diện cho phần chọn mẫu được sử dụng trong việc chọn mẫu có hệ thống của hộ gia đình ở giai đoạn lấy mẫu cuối cùng. Nghịch đảo của nó là khoảng chọn mẫu được áp dụng trong chọn hộ gia đình từ PSU được chọn.

Xem công thức (5), f_d sẽ được thống nhất trong một tầng khi $\frac{a_h}{\sum M_{h\alpha}}$ và b_h không phụ thuộc vào tầng h và do đó, cả hai đều liên tục trên tất cả các tầng trong tầng d . Nên lấy $b_h = 10$, thực hiện cho tất cả các đơn vị chọn mẫu cấp 1, và nếu $\frac{a_h}{\sum M_{h\alpha}}$ có thể được duy trì không đổi, f_d sẽ được thống nhất trong miền d . Để thực hiện về sau, số lượng các PSU được lựa chọn cho tầng h , a_h phải tương ứng với đơn vị đo kích thước của tầng h , mà thực tế là Tổng điều tra dân số năm 2001 là số hộ gia đình cho tầng h . Tuy nhiên, vì a_h phải là một số nguyên và các đơn vị xác định cỡ của tầng cũng khác nhau, xác suất lựa chọn kết quả trên tầng trong tầng d sẽ

không hoàn toàn giống nhau nhưng sẽ không thay đổi nhiều.

Để duy trì một f_d thống nhất trong toàn bộ tầng, khoảng lấy mẫu tương tự có thể được áp dụng trong danh sách tất cả các PSU đã được sắp xếp theo tầng. Điều này ngụ ý rằng việc lựa chọn PSU sẽ không được thực hiện riêng cho mỗi tầng mà sẽ được thực hiện chung cho tất cả các tầng. Quy trình từng bước cho việc duy trì một xác suất chọn thống nhất trong các tầng được nêu dưới đây. Bảng 7 dưới đây sẽ cho thấy xác suất lựa chọn kết quả đồng bộ cho mỗi tên tầng.

Chọn mẫu cho đơn vị chọn mẫu cấp 1

(1) Đối với một tầng d , xác định số lượng PSU được chọn a'_d với $a'_d = \frac{n'_d}{b}$, trong đó b là số lượng hộ gia đình quy định trên mỗi PSU (trong trường hợp này, $b = 10$), n'_d là số hộ gia đình trong tầng d (Bảng 5, cột cuối cùng).

(2) Tính khoảng chọn mẫu:

$$\frac{\sum_{\alpha} M_{h\alpha}}{a'_d} = \frac{N_d}{a_d} \quad (6)$$

(3) Sắp xếp tất cả các PSU trong tầng d của tỉnh, phân loại thành thị/nông thôn, bởi các tầng, bởi giá trị "Pstrong".

(4) Tính giá trị tích lũy của các kích cỡ đo lường (tổng số hộ gia đình dựa trên Tổng điều tra dân số năm 2001), $M_{h\alpha}$ sử dụng danh sách được sắp xếp ở bước (3).

(5) Chọn ngẫu nhiên một số ban đầu (RS) bằng cách lấy một số ngẫu nhiên giữa 0 và 1 và nhân nó với khoảng chọn mẫu trong bước 2. Đơn vị chọn mẫu cấp 1 đầu tiên sẽ là PSU đầu tiên với giá trị tích lũy của $M_{h\alpha}$ chứa giá trị của bắt đầu ngẫu nhiên (RS). Các mẫu PSU tiếp theo sẽ là PSU mà giá trị tích lũy của $M_{h\alpha}$ chứa $RS + S_d$, tiếp theo sẽ là PSU mà giá trị tích lũy $RS + 2 * S_d$; ...

Bảng 7: Tóm tắt các số lượng mẫu thống kê theo tầng

Vùng	Tổng số hộ (N_d)	Ước tính mẫu PSU (a'_d)	Khoảng chọn mẫu (S_d)	Số lượng mẫu PSU thực tế (a_d)	Số lượng mẫu hộ gia đình dự kiến ($hat * n_d$)	Xác suất chọn (f_d)
Barisal	1,648,085	181.77	9066.992	182	1820	0.001104
Chittagong	4,472,548	246.05	18177.35	246	2460	0.000550
Dhaka	8,236,687	369.66	22282.06	370	3700	0.000449
Khulna	3,119,602	210.24	14838.39	210	2100	0.000673
Rajshahi	6,627,797	314.01	21107.21	314	3140	0.000474
Sylhet	1,388,222	178.28	7786.691	178	1780	0.001282

Nguồn: Tính toán của tác giả sử dụng dữ liệu từ cuộc Tổng điều tra dân số năm 2001 tiến hành bởi BBS.

Chọn mẫu hộ gia đình

Vì đơn vị xác định quy mô (ví dụ, tổng số hộ) đã được sử dụng để lựa chọn các PSU là dựa trên Tổng điều tra dân số năm 2001 cách

khá xa giai đoạn Khảo sát LFS 2009-2010, số hộ gia đình được lấy mẫu phải được điều chỉnh cho phù hợp để duy trì sự thống nhất xác suất

lựa chọn trong tầng. Đặc biệt, do các hộ gia đình sẽ được lựa chọn từ một mẫu PSU α với

$$P(h\beta|\alpha) = \frac{b_h}{M_{h\alpha}}$$

và nếu giá trị 2009-2010 của các đơn vị xác định quy mô được ký hiệu là $N_{h\alpha}$, sau đó duy trì cùng một cấp độ hộ gia đình xác suất lựa chọn là:

$$P(h\beta|\alpha) = \frac{b_h}{M_{h\alpha}} = \frac{10}{M_{h\alpha}} = \frac{b'_{h\alpha}}{N_{h\alpha}} \quad (7)$$

và do đó

$$b'_{h\alpha} = \frac{N_{h\alpha}}{M_{h\alpha}} * 10 \quad (8)$$

trong đó $b'_{h\alpha}$ là tổng số lượng các hộ gia đình thực tế được lựa chọn trong PSU α ở tầng h . Điều này cho thấy rằng cần phải có một hoạt động liệt kê tất cả các hộ gia đình ở các PSU được lựa chọn trước khi tiến hành LFS 2009-2010.

3.3. Quyền số điều tra và ước lượng

Quyền số điều tra phải được sử dụng để cho ra các đại lượng ước lượng về tổng thể và đặc điểm thiết kế như các biện pháp phân

tầng, chọn PSU, chọn tầng phải được đưa vào các giá trị ước lượng phương sai và suy luận.

3.3.1. Quyền số điều tra

Các quyền số điều tra cuối cùng là kết quả tính toán của ít nhất ba giai đoạn liên tiếp. Đầu tiên, quyền số cơ sở được tính toán để trung hòa các xác suất lựa chọn không bằng nhau trong việc thiết kế mẫu. Sau đó, các quyền số cơ sở được điều chỉnh để cân bằng tỷ lệ trả lời không đồng đều và nếu dữ liệu có sẵn, quyền số điều chỉnh không trả lời được sẽ tiếp tục điều chỉnh để đảm bảo rằng phân bố quyền số phù hợp với phân bố đã biết từ các nguồn dữ liệu.

Quyền số cơ sở cho mẫu hộ gia đình là nghịch đảo của xác suất lựa chọn của nó. Trong thiết kế mẫu chủ, xác suất lựa chọn là thống nhất trong một tầng và do đó, quyền số cơ sở sẽ không khác nhau trong các tầng. Nói chung

$$w_d^0 = \frac{1}{f_d} \quad (9)$$

Bảng 8 trình bày các quyền số cơ sở của các hộ gia đình được chọn mẫu bằng cách phân tầng:

Bảng 8: Quyền số cơ sở theo tầng

Vùng	Xác suất lựa chọn	Quyền số cơ sở
Barisal	0.001104	905.7971
Chittagong	0.000550	1818.1820
Dhaka	0.000449	2227.1710
Khulna	0.000673	1485.8840
Rajshahi	0.000474	2109.7050
Sylhet	0.001282	780.0312

Nguồn: Tính toán của tác giả sử dụng dữ liệu từ cuộc Tổng điều tra dân số năm 2001 tiến hành bởi BBS.

Điều chỉnh không trả lời sẽ phải được đưa vào quyền số điều tra cuối cùng nếu mức độ của đơn vị không trả lời không thể bỏ qua.

Đơn vị không trả lời xảy ra khi một hộ gia đình đủ điều kiện nhưng không tham gia vào cuộc điều tra. Ví dụ, các hộ gia đình có thể từ chối

tham gia hoặc một đơn vị đủ điều kiện có thể không đồng ý phỏng vấn. Nói chung, việc điều chỉnh không trả lời làm tăng quyền số cơ sở của đơn vị trả lời "tương tự" để bù đắp cho việc không trả lời. Các hình thức phổ biến nhất của việc điều chỉnh quyền số không trả lời là một loại lớp quyền số. Các mẫu đầy đủ của người trả lời và không trả lời được chia thành một số lớp quyền số hoặc các nhóm và các yếu tố điều chỉnh không trả lời được tính cho nhóm (Kalton, 1990) là:

$$w_c^1 = \frac{\sum_{i \in rc} w_{di} + \sum_{j \in rc} w_{dj}}{\sum_{i \in rc} w_{di}} = \frac{\sum_{i \in sc} w_{di}}{\sum_{i \in rc} w_{di}} \quad (10)$$

Trong đó các mẫu số của w_c^1 là tổng quyền số của người được điều tra (chỉ số r) trong quyền số nhóm c trong khi các tử số thêm tổng các quyền số cho người được điều tra và tổng quyền số cho đủ điều kiện không được điều tra (chỉ số m thiếu) trong nhóm c bằng tổng quyền số cho tổng số mẫu đủ điều kiện (chỉ số s) trong nhóm c . Như vậy, việc điều chỉnh quyền số không trả lời w_c^1 là nghịch đảo của tỷ lệ quyền số trả lời trong nhóm c . Lưu ý rằng việc điều chỉnh được áp dụng với các đơn vị đủ điều kiện. Đơn vị lấy mẫu không đủ điều kiện được loại trừ (ví dụ, đơn vị nhà bỏ trống hoặc bị phá hủy và các đơn vị ra khỏi phạm vi được điều tra).

Nhóm quyền số c không cần phù hợp với các ranh giới tầng. Nó có thể cắt qua tầng, nhưng điều quan trọng là các nhóm quyền số sẽ bao quát các hộ "tương tự". Tương tự ở đây được hiểu là "khuynh hướng trả lời của các hộ gia đình tương tự nhau". Nhìn chung, tỷ lệ trả lời trên nhóm quyền số sẽ rất khác nhau. Hơn nữa, có thể có trường hợp mà các quyền số phân bố mẫu sẽ không phù hợp với tính toán tổng thể dự kiến. Khi điều này xảy ra, quyền số điều chỉnh thêm hoặc quyền số điều chỉnh

tổng thể có thể được đưa vào quyền số điều tra cuối cùng để đảm bảo rằng việc phân bố mẫu phù hợp với sự phân bố dân cư. Quyền số điều chỉnh tổng thể được thực hiện tương tự như những quyền số điều chỉnh không trả lời được mô tả trước đó. Phương pháp hiệu chuẩn được sử dụng trong quá trình này. Sử dụng thuật toán phù hợp tỷ lệ lặp đi lặp lại, và được thực hiện trên quyền số không trả lời điều chỉnh như vậy các ước tính quyền số điều tra của một số đặc trưng quan tâm (ví dụ như nhóm tuổi và giới tính) phù hợp với sự phân bố dân số tương ứng.

3.3.2. Ước lượng

Giả sử rằng quyền số điều tra cuối cùng cho hộ gia đình i là w_i hoặc những gì có thể được xem như là số lượng đơn vị dân cư của nhóm mà các hộ gia đình trả lời i đại diện. Sau đó, các ước lượng tổng số dân với đặc trưng của lãi suất sẽ là $\hat{Y} = \sum_{i \in S} w_i y_i$ trong đó y_i là giá trị của biến cho hộ gia đình i .

Ước lượng đơn có nhiều ứng dụng. Ví dụ, nó có thể được áp dụng để ước lượng số dân số có đặc tính cụ thể quan tâm, bằng cách thiết lập $y_i = 1$ nếu hộ gia đình có đặc trưng cụ thể, 0 nếu không.

Để ước lượng trung bình tổng thể, \bar{Y} , các ước lượng tỷ lệ sau đây có thể được sử dụng:

$$\bar{y} = \frac{\sum_{i \in S} w_i y_i}{\sum_{i \in S} y_i} \quad (11)$$

với tổng quyền số điều tra của tất cả các hộ gia đình trả lời, $\sum_{i \in S} w_i$, là một ước lượng cho tổng số các hộ gia đình. Một hình thức tổng quát hơn của các ước lượng tỷ lệ (Kalton, 1983) sẽ là

$$R = \frac{\sum_{i \in S} w_i y_i}{\sum_{i \in S} w_i x_i} \quad (12)$$

Lưu ý rằng với thiết kế mẫu phức tạp như các mẫu chủ, các công cụ mô tả trong (11) và (12) là ước lượng tỷ lệ có liên quan đến tỷ lệ của hai biến ngẫu nhiên và do đó, phải được xem xét một cách cẩn thận trong việc tính toán sai số mẫu.

3.3.3. Ước lượng phương sai

Phương sai là giá trị cần thiết để đánh giá độ chính xác của các cuộc điều tra. Việc thiết kế chọn mẫu, thêm vào kích thước mẫu là rất quan trọng với độ chính xác của ước lượng điều tra. Các gói phần mềm thống kê có mô-đun có thể cho gần đúng các phương sai của các ước tính từ cuộc điều tra phức tạp. Hầu hết các gói phần mềm sử dụng các phương pháp Taylor trong việc tính toán phương sai, mặc dù một số phần mềm cũng cung cấp các phương pháp thay thế trong các hình thức sao chép, tái chọn mẫu hoặc biện pháp khởi động. Nói chung, mỗi phương pháp ước lượng phương sai có ưu điểm và hạn chế riêng.

Lưu ý rằng các ước lượng điều tra tại các phân khu địa lý được dự kiến sẽ có sai số chọn mẫu ở mức độ chấp nhận được. Điều này cũng được dự kiến cho các ước lượng ở cấp quốc gia trên các tầng. Ví dụ, tỷ lệ thất nghiệp ở các cấp độ khu vực thành thị/nông thôn dự kiến sẽ có sai số chọn mẫu chấp nhận được. Điều quan trọng là các sai số chọn mẫu dự toán lớn nên được suy ra để xác nhận các kỳ vọng. Hơn nữa, các sai số chọn mẫu cũng là cần thiết để đánh giá độ tin cậy của các ước tính cấp bộ phận (ví dụ, cấp tỉnh trong trường hợp của

Bangladesh). Ước tính cho bộ phận phụ với đủ cỡ mẫu có thể làm cho các sai số mẫu chấp nhận được. Trong trường hợp của Bangladesh, một số tỉnh vẫn có cỡ mẫu tương đối lớn. Như vậy, mặc dù các bộ phận được thiết lập như là các lĩnh vực thiết kế hoặc tầng rõ ràng, một số ước tính ở mức tỉnh vẫn có thể có sai số chấp nhận được. Tuy nhiên, các ước tính phân tách cấp tỉnh theo thành thị/nông thôn có thể không phải ở tất cả vì không đủ cỡ mẫu.

4. Kết luận

Các kinh nghiệm về quy trình kỹ thuật đã được thực hiện trong sự phát triển của các thiết kế mẫu mới đã được sử dụng cho Khảo sát lực lượng lao động 2009-2010 được tiến hành ở Bangladesh. Các thiết kế mẫu mới giải quyết các điểm yếu được xác định trong các thiết kế trước đó được thông qua vào LFS năm 2005. Một số đề xuất thay đổi như sau: (1) xem xét các tương quan dương trong lớp của tiêu thức quan tâm, thì số hộ gia đình được liệt kê đã giảm từ 40 xuống 10 ở mỗi PSU trong khi số lượng PSU được lựa chọn đã tăng từ 1000 đến 1500. (2) Phương pháp phân bổ mẫu hiệu quả đã được thực hiện để đảm bảo độ tin cậy của ước lượng ở cấp tầng cũng như các đơn vị được phân cắt bởi tầng. (3) Biện pháp phân lớp ngàm đã được giới thiệu để giảm hiệu quả thiết kế. (4) Biện pháp lựa chọn mẫu đảm bảo các xác suất lựa chọn thống nhất cho mỗi bộ phận cũng đã được thông qua để khắc phục các hiệu quả thiết kế lớn được ghi nhận từ LFS năm 2005.

Vân Anh - Thu Hằng (lược dịch)

Nguồn: Developing a Master Sample Design for Household Surveys in Developing Countries: A Case Study in Bangladesh; Dalisay S. Maligalig and Arturo Martinez Jr; 12.07.2013 (<http://surveyinsights.org/?p=2151>)