ภาคผนวก ก แผนการเลือกตัวอย่าง

ภาคผนวก ก จะอธิบายลักษณะสำคัญของแผนการเลือกตัวอย่าง ซึ่งประกอบด้วยการกำหนดกรอบตัวอย่าง (sampling frame) ขนาดตัวอย่างเป้าหมาย (target sample size) การกระจายตัวอย่าง (sample allocation) การนับจดในเขตแจงนับตัวอย่าง (listing in sample clusters) ระดับการนำเสนอผล (choice of domains) ขั้นของการเลือกตัวอย่าง (sampling stages) การจัดชั้นภูมิ (stratification) และการคำนวณค่าถ่วงน้ำหนัก (calculation of sample weights)

การกำหนดแผนการเลือกตัวอย่างสำหรับโครงการสำรวจสถานการณ์เด็กและสตรีในประเทศไทยมีวัตถุประสงค์หลักเพื่อให้ได้ค่าประมาณที่มีความน่าเชื่อถือทางสถิติสำหรับตัวชี้วัดต่าง ๆ ในระดับประเทศ จำแนกตามพื้นที่ในเขตเทศบาลและนอกเขตเทศบาล และในระดับภาค จำนวน 5 ภาค (กรุงเทพมหานคร ภาคกลาง ภาคเหนือ ภาคตะวันออกเฉียงเหนือ และภาคใต้) พื้นที่ในเขตเทศบาลและนอกเขตเทศบาลของแต่ละจังหวัดได้ถูกกำหนดให้เป็นชั้นภูมิสำหรับการเลือกตัวอย่าง ในการออกแบบแผนการเลือกตัวอย่างในครั้งนี้ ได้มีการทบทวนแผนการเลือกตัวอย่างและผลการสำรวจ MICS ปี 2558-2559 ที่ปรากฏในรายงานผลการสำรวจฉบับสมบูรณ์

การสำรวจ MICS พ.ศ. 2562 ได้นำร่องนำเสนอผลการสำรวจระดับจังหวัดจำนวน 17 จังหวัด ซึ่งได้มีการเพิ่มขนาดตัวอย่าง โดย 5 จังหวัดแรก (สงขลา สตูล ปัตตานี ยะลา และนราธิวาส) ได้รับการคัดเลือกเนื่องจากปัญหาความไม่สงบที่มีมายาวนานในพื้นที่ซึ่งประชากรส่วนใหญ่เป็นชาวมุสลิม และปัญหาการเข้าถึงและคุณภาพของสวัสดิการสังคมและการรักษาพยาบาล ส่วน 8 จังหวัดถัดมา (ชัยนาท สระแก้ว แม่ฮ่องสอน ตาก บุรีรัมย์ กาฬสินธุ์ นครพนม และพัทลุง) ได้พิจารณาจากด้านความเสมอภาค โดยใช้เกณฑ์สัดส่วนคนยากจน พ.ศ. 2560 พบว่าเป็นจังหวัดยากจนที่สุดในแต่ละภาค สำหรับจังหวัดที่เหลือ 4 จังหวัด (ราชบุรี กาญจนบุรี ศรีสะเกษ และยโสธร) คือจังหวัดที่เพิ่มขนาดตัวอย่างจากการสำรวจในรอบที่แล้ว จากเกณฑ์ด้านความเสมอภาค ข้อมูลของจังหวัดนำร่องเหล่านี้จะถูกแยกออกมาเพื่อวิเคราะห์และจัดทำรายงานแยกต่างหาก

ในการเลือกตัวอย่างของ MICS กำหนดให้ใช้แผนการเลือกตัวอย่างแบบมีชั้นภูมิหลายขั้น (multi-stage, stratified cluster sampling) กรอบตัวอย่างได้มาจากโครงการสำรวจข้อมูลพื้นฐานของครัวเรือน พ.ศ. 2562 (สพค.62) มีหน่วยตัวอย่างขั้นที่หนึ่งเป็นเขตแจงนับ (Enumeration areas: EAs) ที่กำหนดไว้สำหรับโครงการสำมะโน จากนั้นได้นับจดครัวเรือนทุกหลังในเขตแจงนับตัวอย่าง จึงเลือกครัวเรือนในขั้นที่สอง

ก.1 ขนาดตัวอย่างและการกระจายตัวอย่าง

เนื่องจากขนาดตัวอย่างในภาพรวมของโครงการสำรวจสถานการณ์เด็กและสตรีในประเทศไทย มีส่วนหนึ่งขึ้นอยู่กับระดับการวิเคราะห์และนำเสนอผลในตารางสถิติ ลำดับแรกจึงควรพิจารณาการกระจายตัวของเขตแจงนับและครัวเรือนในกรอบตัวอย่างจาก สพค.62 จำแนกตามภาค จังหวัดนำร่อง พื้นที่ในเขตเทศบาลและนอกเขตเทศบาล ดังที่ปรากฏในตาราง SD.1

ตาราง SD.1: การกระจายเขตแจงนับและครัวเรือนในกรอบตัวอย่าง

การกระจายตัวของเขตแจงนับและครัวเรือน จำแนกตามภาค จังหวัดนำร่อง พื้นที่ในเขตเทศบาลและนอกเขตเทศบาล จาก สพค.62

	จำนวนจังหวัด	จำนวนเขตแจงนับ (สพค.62)			จำนวนครัวเรือน (สพค.62)		
		รวม	ในเขตเทศบาล	นอกเขตเทศบาล	รวม	ในเขตเทศบาล	นอกเขตเทศบาล
ยอดรวม	**77**	**2,985**	**1,530**	**1,455**	**423,526**	**196,136**	**227,336**
ภาค/จังหวัดนำร่อง							
กรุงเทพมหานคร	1	150	150	-	16,174	16,174	-
ภาคกลาง	25	951	450	501	136,463	56,872	79,591
ชัยนาท		33	15	18	4,694	2,096	2,598
สระแก้ว		33	15	18	4,852	1,844	3,008
ราชบุรี		33	15	18	5,326	2,096	3,230
กาญจนบุรี		42	18	24	5,481	1,636	3,845
ภาคเหนือ	17	639	315	324	95,191	42,959	52,232
แม่ฮ่องสอน		39	18	21	4,969	2,354	2,615
ตาก		33	18	15	6,335	2,840	3,495
ภาคตะวันออกเฉียงเหนือ	20	738	366	372	97,167	45,731	51,436
บุรีรัมย์		36	18	18	4,724	2,371	2,353
ศรีสะเกษ		36	18	18	4,467	2,093	2,374
ยโสธร		36	18	18	4,274	2,259	2,015
กาฬสินธุ์		45	21	24	5,370	2,234	3,136
นครพนม		33	18	15	3,972	2,126	1,846
ภาคใต้	14	507	249	258	78,531	34,400	44,131
สงขลา		39	21	18	6,003	2,577	3,426
สตูล		33	15	18	5,265	1,834	3,431
พัทลุง		36	18	18	5,394	2,781	2,613
ปัตตานี		33	15	18	4,728	1,960	2,768
ยะลา		36	18	18	6,605	3,246	3,359
นราธิวาส		33	15	18	5,408	1,925	3,483

สำหรับโครงการสำรวจสถานการณ์เด็กและสตรีในประเทศไทย พ.ศ. 2562 มีขนาดตัวอย่างที่คำนวณได้ทั้งสิ้น จำนวน 40,660 ครัวเรือน โดยตัวชี้วัดหลักที่ใช้ในการคำนวณ คือ ภาวะเตี้ยแคระแกร็น (stunting prevalence) ในเด็กอายุ 0-4 ปี เนื่องจากจะมีการนำเสนอผลการสำรวจในระดับภาค จึงจำเป็นต้องพิจารณาขนาดตัวอย่างขั้นต่ำในแต่ละระดับการนำเสนอ โดยสูตรการคำนวณขนาดตัวอย่างที่ต้องการตามตัวชี้วัดที่กำหนด คือ

$$n = \frac{[4(r)(1-r)(deff)]}{[(RME \times r)^2 (pb)(AveSize)(RR)]},$$

โดยที่

n = ขนาดตัวอย่างที่ต้องใช้ แสดงในรูปของจำนวนครัวเรือนตัวอย่าง

4 = ค่าคงที่เพื่อให้ได้ความเชื่อมั่นที่ระดับ 95%

r = ค่าคาดประมาณของตัวชี้วัด แสดงในรูปของสัดส่วน

deff = ผลกระทบจากการเลือกตัวอย่างที่มีต่อตัวชี้วัด ประมาณค่าจากการสำรวจครั้งที่ผ่านมาหรือใช้ค่ามาตรฐานเท่ากับ 1.5

RME = ขอบเขตของความผิดพลาดของค่า r ที่ระดับความเชื่อมั่น 95% โดยทั่วไปจะมีค่าไม่เกิน 0.12 (ร้อยละ 12) สำหรับค่าประมาณระดับประเทศ

pb = สัดส่วนของประชากรซึ่งใช้เป็นฐานในการคำนวณตัวชี้วัด *r*

AveSize = ขนาดของครัวเรือนเฉลี่ย (จำนวนสมาชิกเฉลี่ยต่อครัวเรือน)

RR = ค่าคาดประมาณของอัตราการตอบ

ในการคำนวณตามสูตร กำหนดให้ *r* (ภาวะเตี้ยแคระแกร็น) เท่ากับร้อยละ 10.5 จากค่าประมาณระดับประเทศของ MICS ปี 2558-2559 *deff* (ผลกระทบจากการเลือกตัวอย่าง) เท่ากับ 5.1 จากค่าประมาณของ MICS ปี 2558-2559 *pb* (ร้อยละของเด็กอายุ 0-4 ปีจากประชากรทั้งสิ้น) เท่ากับร้อยละ 13[1] *AveSize* (ขนาดของครัวเรือนเฉลี่ย) เท่ากับ 3.1 คนต่อครัวเรือน และกำหนดอัตราการตอบเท่ากับร้อยละ 90 จากผลการสำรวจครั้งที่ผ่านมาเช่นกัน และเมื่อนำมาคำนวณตามสูตรจะได้ขนาดตัวอย่างประมาณ 33,300 ครัวเรือน สำหรับค่าประมาณในระดับประเทศ ซึ่งขนาดตัวอย่างดังกล่าวมีค่า RME ประมาณร้อยละ 12

การสำรวจครั้งนี้ยังให้ค่าประมาณระดับภาค (5 ภาค) และจังหวัด (17 จังหวัด จาก 77 จังหวัด) ซึ่งการกระจายของตัวอย่างใน 17 จังหวัดนี้ อยู่ระหว่าง 820 ถึง 1,900 ครัวเรือน ได้ขนาดตัวอย่างสำหรับโครงการสำรวจสถานการณ์เด็กและสตรีในประเทศไทย พ.ศ. 2562 ทั้งสิ้น 40,660 ครัวเรือน โดยกำหนดให้เลือกครัวเรือนตัวอย่าง 20 ครัวเรือน (30 ครัวเรือน สำหรับกรุงเทพมหานคร) ต่อเขตแจงนับ การกำหนดทั้งหมดนี้ได้คำนึงถึงปัจจัยต่าง ๆ ได้แก่ ผลกระทบจากการเลือกตัวอย่าง งบประมาณที่ได้รับ และเวลาที่แต่ละทีมต้องใช้ในการปฏิบัติงานหนึ่งเขตแจงนับ ตาราง SD.2 แสดงการกระจายตัวของเขตแจงนับและครัวเรือนในแต่ละระดับการนำเสนอผล ทั้งนี้ จากผลการสำรวจ MICS ปี 2558-2559 ค่า RME ในระดับจังหวัดมีค่าประมาณร้อยละ 15-23 ยกเว้น กรุงเทพมหานครที่มีค่า RME สูงกว่าเนื่องจากมีขนาดตัวอย่างที่น้อยกว่า

[1] สัดส่วนที่แท้จริงของเด็กอายุต่ำกว่า 5 ปี คิดเป็นร้อยละ 5 ของประชากร แต่เนื่องจากมีการเพิ่มขนาดตัวอย่างของเด็กอายุต่ำกว่า 5 ปี (ดูหัวข้อ ก.3) จึงประมาณค่าสัดส่วนเป็นร้อยละ 13

ตาราง SD.2: การกระจายตัวอย่าง

การกระจายตัวของเขตแจงนับตัวอย่าง (EAs) และครัวเรือนตัวอย่าง จำแนกตามภาค จังหวัดนำร่อง พื้นที่ในเขตเทศบาลและนอกเขตเทศบาล ประเทศไทย พ.ศ. 2562

	เขตแจงนับตัวอย่าง			ครัวเรือนตัวอย่าง		
	รวม	ในเขตเทศบาล	นอกเขตเทศบาล	รวม	ในเขตเทศบาล	นอกเขตเทศบาล
ยอดรวม	**1,958**	**789**	**1,169**	**40,660**	**17,280**	**23,380**
ภาค/จังหวัดนำร่อง						
กรุงเทพมหานคร	150	150	-	4,500	4,500	-
ภาคกลาง	513	227	286	10,260	4,540	5,720
ชัยนาท	95	64	31	1,900	1,280	620
สระแก้ว	81	20	61	1,620	400	1,220
ราชบุรี	95	38	57	1,900	760	1,140
กาญจนบุรี	82	30	52	1,640	600	1,040
ภาคเหนือ	304	95	209	6,080	1,900	4,180
แม่ฮ่องสอน	42	6	36	840	120	720
ตาก	41	11	30	820	220	600
ภาคตะวันออกเฉียงเหนือ	587	172	415	11,740	3,440	8,300
บุรีรัมย์	95	28	67	1,900	560	1,340
ศรีสะเกษ	91	14	77	1,820	280	1,540
ยโสธร	95	25	70	1,900	500	1,400
กาฬสินธุ์	95	49	46	1,900	980	920
นครพนม	73	16	57	1,460	320	1,140
ภาคใต้	404	145	259	8,080	2,900	5,180
สงขลา	92	50	42	1,840	1,000	840
สตูล	46	9	37	920	180	740
พัทลุง	83	42	41	1,660	840	820
ปัตตานี	41	7	34	820	140	680
ยะลา	41	11	30	820	220	600
นราธิวาส	41	8	33	820	160	660

ก.2 การเลือกเขตแจงนับตัวอย่าง และการนับจด

การเลือกหน่วยตัวอย่างขั้นที่หนึ่งใช้กรอบตัวอย่างจากการสำรวจข้อมูลพื้นฐานของครัวเรือน (สพค.) พ.ศ. 2562 ซึ่งดำเนินการระหว่างเดือนตุลาคม-ธันวาคม พ.ศ. 2561 โครงการ สพค. ได้ถูกจัดทำขึ้นทุกปี เพื่อปรับปรุงรายชื่อครัวเรือนในเขตแจงนับตัวอย่างจำนวน 2,985 เขตแจงนับให้เป็นปัจจุบัน (ปรับปรุงจากบัญชีเขตโครงการสำมะโนประชากรและเคหะ พ.ศ. 2553) และใช้เป็นกรอบตัวอย่างสำหรับการสำรวจระดับประเทศในหลาย ๆ โครงการ เช่น การสำรวจภาวะการทำงานของประชากร การสำรวจภาวะเศรษฐกิจและสังคมของครัวเรือน เป็นต้น โครงการ สพค. ใช้แผนการสุ่มตัวอย่างแบบมีชั้นภูมิขั้นเดียว (stratified, single-stage cluster sampling) โดยมีเขตแจงนับ (Enumeration Area: EA) เป็นหน่วยตัวอย่าง และเลือกเขตแจงนับในแต่ละชั้นภูมิ (stratum) ตามความน่าจะเป็นที่เป็นสัดส่วนกับขนาด (Probability Proportional to Size: PPS) โดยหน่วยวัดขนาดคือจำนวนครัวเรือนในแต่ละเขตแจงนับที่ได้จากโครงการสำมะโนประชากรและเคหะ พ.ศ. 2553

สำหรับ 17 จังหวัดนำร่องที่มีการนำเสนอผลระดับจังหวัด พบว่า เขตแจงนับตัวอย่างจากโครงการ สพค. มีจำนวนไม่เพียงพอ จึงต้องทำการเลือกเขตแจงนับเพิ่มเติมจากกรอบตัวอย่างอื่น ซึ่งในที่นี้ได้นำมาจากบัญชีเขตสำมะโนประชากรและเคหะ พ.ศ. 2553 โดยก่อนที่จะทำการเลือกเขตแจงนับเพิ่มเติมต้องนำเขตแจงนับที่ถูกเลือกเป็นตัวอย่างโครงการ สพค. ออกจากกรอบตัวอย่างก่อน จึงทำการเลือกเขตแจงนับเพิ่มเติมในแต่ละชั้นภูมิ (พื้นที่ในเขตเทศบาลและนอกเขตเทศบาลของแต่ละจังหวัด) ด้วยวิธีการสุ่มแบบมีระบบตามความน่าจะเป็นที่เป็นสัดส่วนกับขนาด การนับจดเพิ่มเติมดำเนินการระหว่างเดือนมกราคม-มีนาคม พ.ศ. 2562

ก.3 การเลือกครัวเรือนตัวอย่าง

ทีมนับจดของสำนักงานสถิติจังหวัดและกองบริหารจัดเก็บข้อมูลสถิติ (สำหรับกรุงเทพมหานคร) ได้จัดทำบัญชีรายชื่อครัวเรือนจากงานสนามในแต่ละเขตแจงนับ จากนั้นได้เรียงลำดับครัวเรือนในแต่ละเขตแจงนับเป็นอิสระกันในกลุ่มที่มีเด็กและกลุ่มที่ไม่มีเด็กอายุต่ำกว่า 5 ปี จากนั้นได้เลือกครัวเรือนตัวอย่างจากแต่ละกลุ่มด้วยวิธีการสุ่มแบบมีระบบ เพื่อให้ได้ครัวเรือนตัวอย่างทั้งสิ้น 20 ครัวเรือนในแต่ละเขตแจงนับ (30 ครัวเรือน สำหรับกรุงเทพมหานคร) โดยใช้ซอฟท์แวร์ที่สำนักงานสถิติแห่งชาติเป็นผู้พัฒนาสำหรับการเลือกครัวเรือนอย่างแบบมีระบบ

การสำรวจครั้งนี้ได้ใช้แบบสอบถามสำหรับผู้ชาย ซึ่งจะสัมภาษณ์ครัวเรือนตัวอย่างเพียงครึ่งหนึ่งของจำนวนทั้งสิ้น ทั้งนี้ ซอฟท์แวร์ข้างต้นจะคัดเลือกครัวเรือนตัวอย่างที่ต้องสัมภาษณ์แบบสอบถามสำหรับผู้ชายโดยอัตโนมัติ และผู้ชายอายุ 15-49 ปีทุกคนที่อยู่ในครัวเรือนดังกล่าวเป็นผู้เข้าข่ายที่ต้องสัมภาษณ์

ครัวเรือนในบัญชีรายชื่อของแต่ละเขตแจงนับได้ถูกแบ่งออกเป็น 2 กลุ่มสำหรับการเลือกตัวอย่างขั้นที่สอง คือ กลุ่มครัวเรือนที่มีเด็กอายุต่ำกว่า 5 ปี และกลุ่มครัวเรือนที่ไม่มีเด็กอายุต่ำกว่า 5 ปี การเลือกครัวเรือนตัวอย่างจะดำเนินการแยกจากกันในแต่ละกลุ่ม โดยอัตราการเลือกตัวอย่างจะสูงกว่าในกลุ่มที่มีเด็กอายุต่ำกว่า 5 ปี ด้วยวิธีการนี้ทำให้ได้จำนวนเด็กอายุต่ำกว่า 5 ปีเพิ่มขึ้น ส่งผลให้ความแม่นยำของตัวชี้วัดที่เกี่ยวกับเด็กอายุต่ำกว่า 5 ปีสูงขึ้น

จากจำนวนครัวเรือนตัวอย่าง 20 ครัวเรือน (30 ครัวเรือน สำหรับกรุงเทพมหานคร) ในแต่ละเขตแจงนับ เป้าหมายของจำนวนครัวเรือนที่มีเด็กอายุต่ำกว่า 5 ปี เท่ากับ 10 ครัวเรือน (กรุงเทพมหานคร 15 ครัวเรือน) หากเขตแจงนับใดมีครัวเรือนที่มีเด็กอายุต่ำกว่า 5 ปี มากกว่า 10 ครัวเรือน (กรุงเทพมหานคร 15 ครัวเรือน) จะใช้วิธีการสุ่มแบบมีระบบ เลือกมา 10 ครัวเรือน (กรุงเทพมหานคร 15 ครัวเรือน) และเลือกอีก 10 ครัวเรือน (กรุงเทพมหานคร 15 ครัวเรือน) จากกลุ่มที่ไม่มีเด็กอายุต่ำกว่า 5 ปี สำหรับเขตแจงนับที่มีครัวเรือนที่มีเด็กอายุต่ำกว่า 5 ปี เท่ากับ 10 ครัวเรือนหรือน้อยกว่า (กรุงเทพมหานคร 15 ครัวเรือนหรือน้อยกว่า) ครัวเรือนเหล่านี้จะถูกเลือกทั้งหมด และในเขตแจงนับนั้น จำนวนครัวเรือนที่ไม่มีเด็กอายุต่ำกว่า 5 ปีจะเท่ากับ 20 (30 ครัวเรือน สำหรับกรุงเทพมหานคร) ลบด้วยจำนวนครัวเรือนที่มีเด็กอายุต่ำกว่า 5 ปี

ก.4 การคำนวณค่าถ่วงน้ำหนัก

การเลือกตัวอย่างโครงการ MICS พ.ศ. 2562 ไม่ใช่การใช้ค่าถ่วงน้ำหนักด้วยตนเอง (self-weighting) ด้วยเหตุนี้จึงต้องมีการคำนวณค่าถ่วงน้ำหนักสำหรับใช้ในการวิเคราะห์ข้อมูลต่อไป

องค์ประกอบหลักที่สำคัญของค่าถ่วงน้ำหนัก คือ ส่วนกลับของอัตราการเลือกครัวเรือนตัวอย่างในแต่ละชั้นภูมิ (h) และเขตแจงนับ (i)

$$W_{hi} = \frac{1}{f_{hi}}$$

โดยที่ f_{hi} คือ ความน่าจะเป็นในการเลือกตัวอย่างของเขตแจงนับตัวอย่างที่ i ในชั้นภูมิที่ h ซึ่งก็คือผลคูณของความน่าจะเป็นในการเลือกตัวอย่างของทุก ๆ ขั้น ในแต่ละชั้นภูมิตัวอย่าง ดังนี้

$$f_{hi} = p_{1hi} \times p_{2hi} \times p_{3hi} ,$$

โดยที่ p_{shi} คือ ความน่าจะเป็นในการเลือกตัวอย่างสำหรับการเลือกหน่วยตัวอย่างขั้นที่ s ของเขตแจงนับตัวอย่างที่ i ในชั้นภูมิที่ h จากแผนการเลือกตัวอย่าง ความน่าจะเป็นเหล่านี้คำนวณได้ตามสูตรต่อไปนี้

p_{1hi} = $\frac{n_h \times M_{hi}}{M_h}$,

n_h = จำนวนเขตแจงนับตัวอย่างที่ถูกเลือกในชั้นภูมิที่ h

M_{hi} = จำนวนครัวเรือนจากกรอบตัวอย่างของเขตแจงนับที่ i ในชั้นภูมิที่ h

M_h = จำนวนครัวเรือนทั้งสิ้นจากกรอบตัวอย่างในชั้นภูมิที่ h

p_{2hi} = สัดส่วนของเขตแจงนับที่ถูกนับจดในเขตแจงนับที่ i ในชั้นภูมิที่ h สำหรับ MICS 2562 กำหนดให้ $p_{2hi} = 1$ ในทุกเขตแจงนับ เนื่องจากไม่มีการแบ่งส่วนย่อยในเขตแจงนับ

p_{3hi} = ความน่าจะเป็นขั้นสุดท้ายในการเลือกครัวเรือนจากกรอบตัวอย่างในเขตแจงนับที่ i เป็นอิสระกันในกลุ่มที่มีเด็กและกลุ่มที่ไม่มีเด็กอายุต่ำกว่า 5 ปี ตามรายละเอียดด้านล่าง

ความน่าจะเป็นในการเลือกตัวอย่างขั้นสุดท้ายในแต่ละเขตแจงนับของกลุ่มครัวเรือนที่มีเด็กอายุต่ำกว่า 5 ปี จะแตกต่างจากกลุ่มที่ไม่มีเด็ก ดังนั้น จึงคำนวณค่าถ่วงน้ำหนักของแต่ละกลุ่มครัวเรือนดังกล่าวในแต่ละเขตแจงนับแยกจากกัน

จากแผนการเลือกตัวอย่างแบบมีชั้นภูมิสองขั้น ความน่าจะเป็นในการเลือกครัวเรือนตัวอย่างที่มีเด็กอายุต่ำกว่า 5 ปี ภายในเขตแจงนับ สามารถคำนวณได้ตามสูตร

$$f_{hi(wc)} = \frac{n_h \times M_{hi}}{M_h} \times p_{2hi} \times \frac{m_{hi(wc)}}{M'_{hi(wc)}} ,$$

โดยที่

$f_{hi(wc)}$ = ความน่าจะเป็นในการเลือกครัวเรือนตัวอย่างที่มีเด็กอายุต่ำกว่า 5 ปี ของเขตแจงนับที่ *i* ในชั้นภูมิที่ *h*

n_h, M_{hi}, M_h และ p_{2hi} มีความหมายเช่นเดียวกับในสูตรก่อนหน้านี้

$m_{hi(wc)}$ = จำนวนครัวเรือนตัวอย่างที่มีเด็กอายุต่ำกว่า 5 ปี ที่ถูกเลือกของเขตแจงนับที่ *i* ในชั้นภูมิที่ *h*

$M'_{hi(wc)}$ = จำนวนครัวเรือนตัวอย่างที่มีเด็กอายุต่ำกว่า 5 ปีทั้งสิ้น ที่ถูกนับจดของเขตแจงนับที่ *i* ในชั้นภูมิที่ *h*

ในทำนองเดียวกัน ความน่าจะเป็นในการเลือกครัวเรือนตัวอย่างที่ไม่มีเด็กอายุต่ำกว่า 5 ปี สามารถคำนวณได้ตามสูตร

$$f_{hi(woc)} = \frac{n_h \times M_{hi}}{M_h} \times p_{2hi} \times \frac{m_{hi(woc)}}{M'_{hi(woc)}},$$

โดยที่

$f_{hi(woc)}$ = ความน่าจะเป็นในการเลือกครัวเรือนตัวอย่างที่ไม่มีเด็กอายุต่ำกว่า 5 ปี ของเขตแจงนับที่ *i* ในชั้นภูมิที่ *h*

$m_{hi(woc)}$ = จำนวนครัวเรือนตัวอย่างที่ไม่มีเด็กอายุต่ำกว่า 5 ปี ที่ถูกเลือกของเขตแจงนับที่ *i* ในชั้นภูมิที่ *h*

$M'_{hi(woc)}$ = จำนวนครัวเรือนตัวอย่างที่มีเด็กอายุต่ำกว่า 5 ปีทั้งสิ้น ที่ถูกนับจดของเขตแจงนับที่ *i* ในชั้นภูมิที่ *h*

เนื่องจากจำนวนครัวเรือนทั้งสิ้นในแต่ละเขตแจงนับจากกรอบตัวอย่างที่ใช้ในการเลือกหน่วยตัวอย่างขั้นที่หนึ่ง แตกต่างจากจำนวนครัวเรือนที่มีและไม่มีเด็กอายุต่ำกว่า 5 ปี ในแต่ละเขตแจงนับที่ได้จากการนับจด จึงต้องคำนวณความน่าจะเป็นในการเลือกครัวเรือนที่มีและไม่มีเด็กอายุต่ำกว่า 5 ปี ในแต่ละเขตแจงนับ

องค์ประกอบสุดท้ายในการคำนวณค่าถ่วงน้ำหนัก คือ การปรับค่าถ่วงน้ำหนักโดยการใช้อัตราการไม่ตอบทั้งในระดับครัวเรือน และระดับบุคคล โดยการปรับค่าสำหรับครัวเรือนที่ไม่ตอบในแต่ละเขตแจงนับ คือ

$$\frac{1}{RR_h}$$

โดยที่ RR_h คือ อัตราการตอบของครัวเรือนตัวอย่างในชั้นภูมิที่ *h* คำนวณจากสัดส่วนของจำนวนครัวเรือนที่สัมภาษณ์ได้ในชั้นภูมิที่ *h* หารด้วย จำนวนครัวเรือนตัวอย่างที่มีคนอยู่อาศัยระหว่างคาบการปฏิบัติงานสนามในชั้นภูมิที่ *h*

ในทำนองเดียวกัน การปรับค่าถ่วงน้ำหนักโดยการใช้อัตราการไม่ตอบในระดับบุคคล (ผู้หญิง ผู้ชาย และเด็กอายุต่ำกว่า 5 ปี) สำหรับแต่ละชั้นภูมิ คือ

$$\frac{1}{RR_{qh}}$$

โดยที่ RR_{qh} คือ อัตราการตอบของแบบสอบถามรายบุคคลในชั้นภูมิที่ *h* คำนวณจากสัดส่วนของผู้เข้าข่าย (ผู้หญิง ผู้ชาย หรือเด็กอายุต่ำกว่า 5 ปี) ในชั้นภูมิที่ *h* ที่สัมภาษณ์ได้สมบูรณ์

หลังจากการปฏิบัติงานสนามเสร็จสิ้น ได้มีการคำนวณอัตราการตอบแบบสอบถาม (response rate) ในแต่ละชั้นภูมิ จำแนกตามกลุ่มครัวเรือนที่มีและไม่มีเด็กอายุต่ำกว่า 5 ปี แล้วนำค่าดังกล่าวไปปรับค่าถ่วงน้ำหนักสำหรับแต่ละเขตแจงนับ อัตราการตอบได้แสดงไว้ในตาราง SR.1.1 ของรายงานผลการสำรวจฉบับนี้

จากนั้นนำตัวแปรที่ปรับอัตราการไม่ตอบสำหรับแบบสอบถามสำหรับผู้หญิง และเด็กอายุต่ำกว่า 5 ปี ไปใช้กับค่าถ่วงน้ำหนักในระดับครัวเรือนที่ปรับแล้ว โดยจำนวนผู้เข้าข่ายทั้งผู้หญิง และเด็กอายุต่ำกว่า 5 ปี นำมาจากรายชื่อสมาชิกในครัวเรือนในแบบสอบถามสำหรับครัวเรือนที่สัมภาษณ์ได้สมบูรณ์

สำหรับค่าถ่วงน้ำหนักของแบบสอบถามสำหรับผู้ชาย สามารถคำนวณได้ในลักษณะเดียวกัน โดยตัวแปรที่ปรับอัตราการไม่ตอบมีตัวตั้งเป็นจำนวนผู้ชายที่เข้าข่ายจากรายชื่อสมาชิกในครัวเรือนของครัวเรือนตัวอย่างทั้งสิ้นในชั้นภูมินั้น ๆ และจำนวนแบบสอบถามสำหรับผู้ชายที่สมบูรณ์ในแต่ละชั้นภูมิ คิดจากการเลือกตัวอย่างย่อยร้อยละ 50 ของครัวเรือน ดังนั้น ตัวแปรที่ปรับค่าถ่วงน้ำหนักได้รวมตัวแปรถ่วงน้ำหนักจากการเลือกตัวอย่างย่อยซึ่งมีค่าเท่ากับ 2 ไว้แล้ว เพิ่มเติมจากการปรับการไม่ตอบแบบสอบถามสำหรับผู้ชาย

กรณีแบบสอบถามสำหรับเด็กอายุ 5-14 ปี ซึ่งในแต่ละครัวเรือนมีการเลือกเด็กในกลุ่มอายุนี้มาหนึ่งคนจากรายชื่อของสมาชิกในครัวเรือน จึงมีการปรับค่าถ่วงน้ำหนักครัวเรือนสำหรับเด็กอายุ 5-14 ปี จากอัตราการตอบแบบสอบถามนี้ในระดับชั้นภูมิ เมื่อค่าถ่วงน้ำหนักครัวเรือนที่ปรับแล้วได้ทำการนอร์มัลไลซ์ (normalized) ดังคำอธิบายในย่อหน้าถัดไป จึงคูณด้วยจำนวนเด็กอายุ 5-14 ปี ที่ปรากฏในรายชื่อของสมาชิกในครัวเรือน ดังนั้นค่าถ่วงน้ำหนักของเด็กอายุ 5-14 ปี จะแปรผันไปตามครัวเรือนตัวอย่าง การถ่วงน้ำหนักข้อมูลของเด็กอายุ 5-14 ปี ได้ดำเนินการในโปรแกรมการออกตารางสำหรับตารางสถิติที่เกี่ยวข้อง

ค่าถ่วงน้ำหนักตามแผนการเลือกตัวอย่าง (full or raw weights) สำหรับครัวเรือน คำนวณโดยการคูณส่วนกลับของความน่าจะเป็นในการเลือกกับตัวแปรที่ปรับการไม่ตอบในแต่ละเขตแจงนับ จำแนกตามกลุ่มครัวเรือนที่มีและไม่มีเด็กอายุต่ำกว่า 5 ปี จากนั้นนำค่าถ่วงน้ำหนักที่ได้มาทำให้เป็นมาตรฐานหรือการนอร์มัลไลซ์ (normalized) เพื่อทำให้ผลรวมค่าถ่วงน้ำหนักของครัวเรือนที่สัมภาษณ์ได้มีค่าเท่ากับจำนวนครัวเรือนตัวอย่างที่สัมภาษณ์ได้ทั่วประเทศ ซึ่งการนอร์มัลไลซ์นี้เป็นการนำค่าถ่วงน้ำหนักที่ปรับอัตราการไม่ตอบแล้วหารด้วยค่าเฉลี่ยของค่าถ่วงน้ำหนักของครัวเรือนตัวอย่างที่สัมภาษณ์ได้ทั่วประเทศ ทำได้โดยการคูณค่าถ่วงน้ำหนักกับค่าคงที่ซึ่งเท่ากับจำนวนครัวเรือนทั่วประเทศที่ยังไม่ได้ถ่วงน้ำหนักหารด้วยจำนวนครัวเรือนทั้งสิ้นที่ถ่วงน้ำหนัก (ใช้ค่าถ่วงน้ำหนักที่ปรับการไม่ตอบ) สำหรับค่าถ่วงน้ำหนักของผู้หญิง ผู้ชาย และเด็กอายุต่ำกว่า 5 ปี ได้ทำการนอร์มัลไลซ์ด้วยวิธีการเช่นเดียวกันนี้ ค่าถ่วงน้ำหนักที่ทำการนอร์มัลไลซ์แล้วของครัวเรือน มีค่าระหว่าง 0.009261 - 17.872195 ใน 1,958 เขตแจงนับตัวอย่าง

หลังจากนั้นจึงนำค่าถ่วงน้ำหนักไปเพิ่มต่อท้ายในชุดข้อมูลทุกชุด และทำการวิเคราะห์ข้อมูลครัวเรือน ผู้หญิง ผู้ชาย เด็กอายุต่ำกว่า 5 ปี และเด็กอายุ 5-14 ปี โดยใช้ค่าถ่วงน้ำหนักเหล่านี้