



รายงานวิจัยฉบับสมบูรณ์

โครงการ “การจัดทำฐานข้อมูลแบบตัวอย่างซ้ำจากข้อมูลภาวะการทำงาน
ของประชากร”

Constructing a Panel Data from the Labor Force Survey of
Thailand

เสนอต่อ

สำนักงานกองทุนสนับสนุนการวิจัย (สกว.)

โดย

นางสาวภัทรพรรณ อดทน

สถาบันวิจัยเพื่อการประเมินและออกแบบนโยบาย

1 กุมภาพันธ์ 2561 (สิ้นสุดโครงการ)

รายงานวิจัยฉบับสมบูรณ์

โครงการ “การจัดทำฐานข้อมูลแบบตัวอย่างซ้ำจากข้อมูลภาวะการทำงาน
ของประชากร”

Constructing a Panel Data from the Labor Force Survey of
Thailand

คณะผู้วิจัย

นางสาวภัทรพรรณ อดทน สถาบันวิจัยเพื่อการประเมินและออกแบบนโยบาย

สนับสนุนโดยสำนักงานกองทุนสนับสนุนการวิจัย (สกว.)

(ความเห็นในรายงานนี้เป็นของผู้วิจัย สกว. ไม่จำเป็นต้องเห็นด้วยเสมอไป)

กิตติกรรมประกาศ

งานวิจัยชิ้นนี้ได้รับเงินทุนสนับสนุนจากสำนักงานกองทุนสนับสนุนการวิจัย (สกว.) ภายใต้ชุดโครงการพัฒนาองค์ความรู้เศรษฐกิจและสังคมของครัวเรือนไทย คณะผู้วิจัยขอขอบคุณ คุณรจนา เนตรแสงทิพย์ รองปลัดกระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม ที่ได้ให้ข้อคิดเห็นเพื่อเป็นแนวทางในการดำเนินงานของโครงการ และขอขอบคุณผู้เชี่ยวชาญจากสำนักงานสถิติแห่งชาติ คุณอานนท์ จันทวิช ผู้ตรวจราชการกรม, คุณน้ำผึ้ง เชิดชูพงษ์ ผู้อำนวยการกองเศรษฐกิจและสังคม, คุณบงกช วิบูลย์ธนานันต์ ผู้อำนวยการกองนโยบายและวิชาการสถิติ, คุณจิรดา ชัยเดชอัครกุล ผู้อำนวยการกลุ่มสถิติแรงงาน, คุณหทัยชนก พรหมเจริญ ผู้อำนวยการกลุ่มระเบียบวิธีสถิติ, คุณพรพรรณ แก้วศรีงาม ผู้อำนวยการกลุ่มคลังข้อมูล และคุณวรัญญา สุขวงศ์ นักวิชาการสถิติชำนาญการ ที่ได้ให้ข้อเสนอแนะเกี่ยวกับระเบียบวิธีการวิจัย การเผยแพร่ฐานข้อมูลและการจัดทำฐานข้อมูล อีกทั้งคณะผู้วิจัยขอขอบคุณ ดร.นงา วะสี หัวหน้ากลุ่มงานวิจัย สถาบันวิจัยเศรษฐกิจป๋วย อึ๊งภากรณ์ ธนาคารแห่งประเทศไทย ที่ให้คำปรึกษาแนะนำการจัดทำโครงการจนประสบผลสำเร็จลุล่วงไปได้ด้วยดี

คณะนักวิจัย

บทสรุปผู้บริหาร

โครงการวิจัยนี้มีวัตถุประสงค์เพื่อจัดทำฐานข้อมูลรายบุคคลแบบตัวอย่างซ้ำจากข้อมูลส่วน Out Rotation Group (ORG) ของข้อมูลการสำรวจภาวะการทำงานของประชากร (Labor Force Survey: LFS) ซึ่งจะเป็นประโยชน์ต่อการศึกษาวิจัยในประเด็นสำคัญต่าง ๆ เช่น การย้ายถิ่นของประชากร การเปลี่ยนงานของแรงงาน พลวัตของการจ้างงาน ผลกระทบของการเปลี่ยนแปลงระบบการกำหนดค่าจ้างขั้นต่ำ เป็นต้น

นักวิจัยได้ดำเนินการศึกษาแผนการเลือกตัวอย่างแบบ rotation group ของโครงการสำรวจภาวะการทำงานของประชากรไทย ในปี พ.ศ. 2545 – 2556 ซึ่งแผนการเลือกตัวอย่างจะถูกแบ่งออกเป็น 2 ช่วง คือ ช่วงปี พ.ศ. 2545 – 2554 และช่วงปี พ.ศ. 2555 – 2556 โดยที่ในแต่ละ rotation plan ครึ่งเรือนจะถูกสัมภาษณ์ซ้ำมากที่สุด 4 ครั้งในรอบ 2 ปีเท่านั้น ทั้งนี้จึงขึ้นอยู่กับผู้ใช้งานในการพิจารณาเลือกใช้กลุ่มตัวอย่างจากข้อมูลชุดนี้

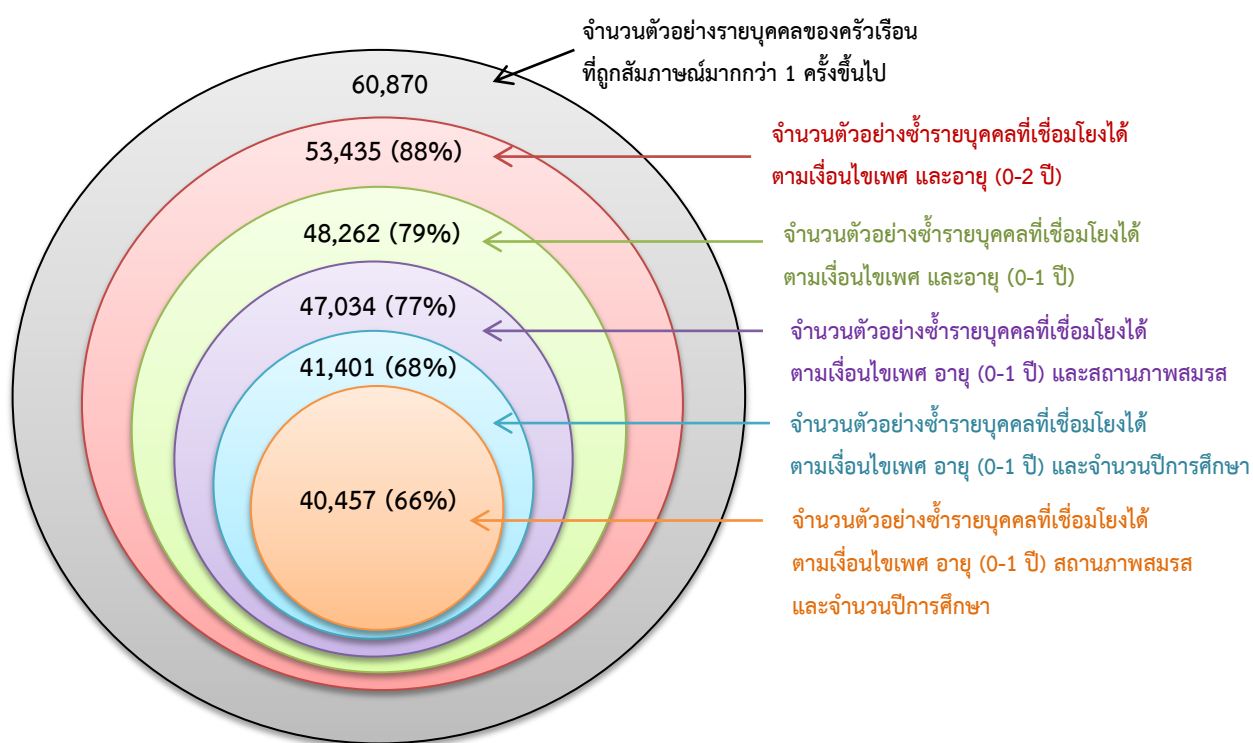
เมื่อได้จัดกลุ่มตัวอย่างเป็นชุดตาม rotation plan พบว่ามีจำนวนครัวเรือนเฉลี่ยชุดละ 18,072 ครัวเรือน ซึ่งครัวเรือนที่ถูกสัมภาษณ์อย่างน้อย 1 ครั้ง มีจำนวนเฉลี่ย 17,179 ครัวเรือน คิดเป็นร้อยละ 95 ของจำนวนครัวเรือนทั้งหมด โดยที่ครัวเรือนที่ถูกสัมภาษณ์อย่างน้อย 1 ครั้ง มีจำนวนตัวอย่างรายบุคคลเฉลี่ยชุดละ 60,870 คน

ในการดำเนินการเชื่อมโยงบุคคลตัวอย่างภายใต้สมมติฐานที่ว่า กลุ่มตัวอย่างส่วนใหญ่ใช้ลำดับสมาชิกเดิมในรอบการสำรวจติดกัน จึงใช้ caseid เป็นตัวแปรในการเชื่อมโยงบุคคลข้ามรอบการสำรวจ จากนั้นตรวจสอบเพศ อายุ ระดับการศึกษา และสถานภาพสมรส ว่าสอดคล้องตามเกณฑ์หรือไม่ ส่งผลให้กระบวนการนี้สามารถแบ่งกลุ่มตัวอย่างเป็น 2 กลุ่มใหญ่ ๆ คือ กลุ่มที่สามารถเชื่อมโยงข้อมูลได้ด้วย caseid เดิม กับกลุ่มที่ไม่สามารถใช้ caseid เดิมในการเชื่อมโยงข้อมูลได้

สำหรับกลุ่มที่ไม่สามารถเชื่อมโยงด้วย caseid เดิมได้นั้น จะถูกนำมาตรวจสอบเพศ และอายุ เพื่อเชื่อมโยงว่าเป็นสมาชิกคนใดในบ้านนั้น ๆ ในรอบถัดไปของการถูกสัมภาษณ์ และสร้าง caseid ให้ใหม่ หลังจากนั้นจะนำข้อมูลจากทั้ง 2 ส่วนมารวมกัน โดยมีจำนวนตัวอย่างซ้ำรายบุคคลที่ตรวจสอบจาก เพศ อายุ ระดับการศึกษา และสถานภาพสมรส เฉลี่ยชุดละ 40,457 คน คิดเป็นร้อยละ 66 ของจำนวนคนทั้งหมด และเมื่อพิจารณาเพศ อายุ และระดับการศึกษา จะมีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ยชุดละ 41,401 คน คิดเป็นร้อยละ 68 ของจำนวนคนทั้งหมด แต่หากพิจารณาเพศ อายุ และสถานภาพสมรส จะมีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ยชุดละ 47,034 คน คิดเป็นร้อยละ

77 ของจำนวนคนทั้งหมด ในขณะที่หากตรวจสอบแค่เพศ และอายุ โดยที่คนเดียวกันต้องมีอายุไม่เกิน 1 ปี จะทำให้มีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ยชุดละ 48,262 คน คิดเป็นร้อยละ 79 ของจำนวนคนทั้งหมด และเมื่อตรวจสอบเพศ และอายุ โดยที่คนเดียวกันต้องมีอายุไม่เกิน 2 ปี จะทำให้มีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ยชุดละ 53,435 คน คิดเป็นร้อยละ 88 ของจำนวนคนทั้งหมด (ตามแผนภาพด้านล่าง)

จากสมมติฐานข้อกำหนดในการเชื่อมโยงบุคคลตัวอย่างซ้ำที่ใช้เพศและอายุเป็นหลักนั้น (กรณีที่ไม่สามารถใช้ caseid ในการเชื่อมโยงได้) อาจเกิดข้อผิดพลาดในกรณีที่สมาชิกในครัวเรือนเดียวกัน มีเพศเดียวกัน มีอายุเท่ากัน และอยู่ในรอบเดียวกันที่ถูกสัมภาษณ์ (เช่น ครัวเรือนที่มีสมาชิกเป็นฝาแฝด) ทำให้การเชื่อมโยงสมาชิกในรอบถัดไปเกิดการสลับที่ได้ โดยมีโอกาสเชื่อมโยงผิดพลาดเฉลี่ยร้อยละ 0.4 ของจำนวนตัวอย่างซ้ำที่ใช้เพศและอายุในการเชื่อมโยง



แผนภาพสรุปจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ยต่อ 1 rotation plan

บทคัดย่อ

โครงการการจัดทำฐานข้อมูลแบบตัวอย่างซ้ำจากข้อมูลภาวะการทำงานของประชากร (Constructing a Panel Data from the Labor Force Survey of Thailand) จะนำข้อมูลการสำรวจภาวะการทำงานของประชากร (Labor Force Survey: LFS) นับตั้งแต่ปี พ.ศ. 2545 – 2556 มาสร้างข้อมูลตัวอย่างซ้ำด้านแรงงานของประชากรไทย จากข้อมูลส่วน Outgoing Rotation Group (ORG) ซึ่งจะเป็นประโยชน์ต่อการศึกษาวิจัยในประเด็นสำคัญต่าง ๆ เช่น การย้ายถิ่นของประชากร การเปลี่ยนงานของแรงงาน พลวัตของการจ้างงาน ผลกระทบของการเปลี่ยนแปลงระบบการกำหนดค่าจ้างขั้นต่ำ เป็นต้น โดยโครงการฯ จะเผยแพร่ข้อมูลส่วนที่เป็น Matching ID ที่สร้างขึ้น พร้อมทั้ง STATA code และคู่มือการสร้างข้อมูลตัวอย่างซ้ำในเว็บไซต์เพื่อให้ นักวิจัยไทยได้ใช้อย่างทั่วถึง

ฐานข้อมูลที่ได้จากกระบวนการวิจัยของโครงการนี้เป็นข้อมูลรายบุคคลแบบตัวอย่างซ้ำซึ่ง นักวิจัยด้านเศรษฐศาสตร์แรงงานหรือนักวิจัยด้านอื่น ๆ ที่เกี่ยวข้องกับการใช้ข้อมูลตัวอย่างซ้ำสามารถนำไปใช้ในการศึกษาวิจัยต่อไปได้ ซึ่งถือเป็นการลดต้นทุนในการสร้างข้อมูล

Abstract

This research project attempts to construct a panel data from the series of Thailand Labor Force Survey data during the years 2002 – 2013. Results from this project could be useful for future research, e.g. the study on migration and job change, the dynamics of employment, and the impact of minimum wage system. STATA code, matching ID, and user manual are publicly available on web site for easy accessed by the researchers.

Database derived from this research project are replicated individual data. These help the labor economist and other researchers to reduce the cost of data management.

สารบัญ

เรื่อง	หน้า
กิตติกรรมประกาศ	i
บทสรุปผู้บริหาร	ii
บทคัดย่อ	iv
สารบัญ	vi
สารบัญภาพ	vii
สารบัญตาราง	xiv
นิยามศัพท์	xvi
บทที่ 1 บทนำ	1
บทที่ 2 ทบทวนวรรณกรรม	3
บทที่ 3 กระบวนการจัดทำฐานข้อมูลรายบุคคลแบบตัวอย่างซ้ำ จากข้อมูลส่วน Out Rotation Group (ORG)	7
บทที่ 4 ผลการวิจัย	19
บทที่ 5 บทสรุป	41
เอกสารอ้างอิง	46
ภาคผนวก ก. สถิติเบื้องต้นจากข้อมูลรายบุคคลแบบตัวอย่างซ้ำจากข้อมูลภาวะการทำงาน	47
ภาคผนวก ข. ชุดคำสั่งในการจัดการข้อมูลรายบุคคลแบบตัวอย่างซ้ำจากข้อมูลภาวะการทำงาน	113
ภาคผนวก ค. สรุปรายงานการประชุมกับผู้เชี่ยวชาญจากสำนักงานสถิติแห่งชาติ	172
คู่มือการใช้โค้ดคำสั่ง ORG	183

สารบัญภาพ

รูปที่		หน้า
3.1	การเลือกตัวอย่างแบบ 4 Rotation groups และ 2-2-2 pattern ที่ใช้ในปี พ.ศ. 2545 – 2554	8
3.2	การเลือกตัวอย่างแบบ 4 Rotation groups และ 2-2-2 pattern ที่ใช้ในปี พ.ศ. 2555 – 2564	9
3.3	การเชื่อมโยงข้อมูลตัวอย่างข้าราชการบุคคลจากข้อมูลการสำรวจภาวะการทำงานของประชากร	11
3.4	เงื่อนไขของการเปลี่ยนสถานภาพสมรส	13
3.5	กระบวนการศึกษาข้อมูลกลุ่มตัวอย่างข้าราชการบุคคลจากข้อมูลการสำรวจภาวะการทำงานของประชากร	18
5.1	การเชื่อมโยงข้อมูลตัวอย่างข้าราชการบุคคลจากข้อมูลการสำรวจภาวะการทำงานของประชากร	41
5.2	จำนวนข้อมูลตัวอย่างข้าราชการครัวเรือนเฉลี่ยต่อ 1 rotation plan	43
5.3	จำนวนข้อมูลตัวอย่างข้าราชการบุคคลเฉลี่ยต่อ 1 rotation plan	44
1	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่แบ่งตามเพศ	48
2	อายุเฉลี่ยของกลุ่มตัวอย่างจากข้อมูล LFS และ ORG	49
3	ค่าจ้างเฉลี่ยรายชั่วโมงของข้อมูล LFS และ ORG	50
4	ค่าจ้างเฉลี่ยรายสัปดาห์จากข้อมูล LFS และ ORG	50
5	ค่าจ้างเฉลี่ยรายเดือนจากข้อมูล LFS และ ORG	51
6	ค่าโบนัสเฉลี่ยรายปีจากข้อมูล LFS และ ORG	51
7	ค่าโอทีเฉลี่ยรายเดือนจากข้อมูล LFS และ ORG	52

สารบัญภาพ

รูปที่		หน้า
8	ร้อยละโดยเฉลี่ยตั้งแต่ปี 2002-2013 ของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS แบ่งตามระดับการศึกษา	53
9	ร้อยละโดยเฉลี่ยตั้งแต่ปี 2002-2013 ของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG แบ่งตามระดับการศึกษา	54
10	ร้อยละโดยเฉลี่ยตั้งแต่ปี 2002-2013 ของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 2 waves แบ่งตามระดับการศึกษา	55
11	ร้อยละโดยเฉลี่ยตั้งแต่ปี 2002-2013 ของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 3 waves แบ่งตามระดับการศึกษา	56
12	ร้อยละโดยเฉลี่ยตั้งแต่ปี 2002-2013 ของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 4 waves แบ่งตามระดับการศึกษา	57
13	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่ไม่มีการศึกษา	58
14	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาต่ำกว่าชั้นประถมศึกษา	58
15	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาชั้นประถมศึกษา	59
16	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาชั้นมัธยมศึกษาตอนต้น	60
17	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาชั้นมัธยมศึกษาตอนปลาย	60

สารบัญภาพ

รูปที่		หน้า
18	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาชั้นอาชีวศึกษาระดับ ปวช. (เทียบเท่ามัธยมศึกษาตอนปลายสาย สามัญ)	61
19	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษายสายศึกษาศาสตร์ (เทียบเท่ามัธยมศึกษาตอนปลายสายสามัญ)	62
20	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาชั้นอนุปริญญา (สายสามัญ)	62
21	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาชั้นอนุปริญญา (สายอาชีวศึกษา ระดับ ปวส.)	63
22	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาชั้นอนุปริญญา (สายศึกษาศาสตร์)	64
23	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาระดับปริญญาตรี (สายวิชาการ)	64
24	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาระดับปริญญาตรี (สายอาชีวศึกษา)	65
25	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาระดับปริญญาตรี (สายศึกษาศาสตร์)	66
26	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาระดับปริญญาโท	66
27	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาระดับปริญญาเอก	67
28	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG แบ่งตามกลุ่มอาชีพ	68

สารบัญภาพ

รูปที่		หน้า
29	ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS ตั้งแต่ปี 2002-2013 แบ่งตามกลุ่มอาชีพ	69
30	ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG ตั้งแต่ปี 2002-2013 แบ่งตามกลุ่มอาชีพ	70
31	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 2 waves ตั้งแต่ปี 2002-2013 แบ่งตามกลุ่มอาชีพ	71
32	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 3 waves ตั้งแต่ปี 2002-2013 แบ่งตามกลุ่มอาชีพ	72
33	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 4 waves ตั้งแต่ปี 2002-2013 แบ่งตามกลุ่มอาชีพ	73
34	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่ประกอบอาชีพ นายจ้าง	74
35	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่ประกอบธุรกิจส่วนตัวโดยไม่มีลูกจ้าง	74
36	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่ช่วยธุรกิจครัวเรือนโดยไม่ได้รับค่าจ้าง	75
37	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่ประกอบอาชีพลูกจ้างรัฐบาล	76
38	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่ประกอบอาชีพลูกจ้างรัฐวิสาหกิจ	76
39	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่ประกอบอาชีพลูกจ้างเอกชน	77

สารบัญภาพ

รูปที่		หน้า
40	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่ประกอบอาชีพโดยการรวมกลุ่ม	78
41	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG แบ่งตามความสัมพันธ์กับหัวหน้าครัวเรือน	79
42	ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS ตั้งแต่ปี 2002-2013 แบ่งตามความสัมพันธ์กับหัวหน้าครัวเรือน	80
43	ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG ตั้งแต่ปี 2002-2013 แบ่งตามความสัมพันธ์กับหัวหน้าครัวเรือน	81
44	ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 2 waves ตั้งแต่ปี 2002-2013 แบ่งตามความสัมพันธ์กับหัวหน้าครัวเรือน	82
45	ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 3 waves ตั้งแต่ปี 2002-2013 แบ่งตามความสัมพันธ์กับหัวหน้าครัวเรือน	83
46	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 4 waves ตั้งแต่ปี 2002-2013 แบ่งตามความสัมพันธ์กับหัวหน้าครัวเรือน	84
47	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่เป็นหัวหน้า ครัวเรือน	85
48	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่เป็นสามีหรือภรรยา	85
49	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่เป็นบุตรที่ยังไม่ สมรส	86
50	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่เป็นบุตรที่สมรสแล้ว	87
51	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่เป็นบุตรเขยหรือ บุตรสะใภ้	87
52	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่เป็นบุตรของบุตร	88

สารบัญภาพ

รูปที่		หน้า
53	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่เป็นพ่อ หรือแม่ หรือพ่อตา หรือแม่ยาย	89
54	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่เป็นญาติอื่น ๆ	89
55	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่เป็นผู้อาศัย และคนรับใช้	90
56	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG แบ่งตามสถานภาพ	91
57	ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS ตั้งแต่ปี 2002-2013 แบ่งตามสถานภาพ	92
58	ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG ตั้งแต่ปี 2002-2013 แบ่งตามสถานภาพ	93
59	ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 2 waves ตั้งแต่ปี 2002-2013 แบ่งตามสถานภาพ	94
60	ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 3 waves ตั้งแต่ปี 2002-2013 แบ่งตามสถานภาพ	95
61	ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 4 waves ตั้งแต่ปี 2002-2013 แบ่งตามสถานภาพ	96
62	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG สถานภาพโสด	97
63	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG สถานภาพสมรส	97
64	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG สถานภาพหย่า	98
65	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG สถานภาพแยกกันอยู่	98
66	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่เคยสมรสแต่ไม่ทราบสถานภาพ	99

สารบัญภาพ

รูปที่		หน้า
67	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG แบ่งตามเขตที่อยู่อาศัย	100
68	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่อยู่อาศัยในเขตเทศบาล	101
69	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่อยู่อาศัยนอกเขตเทศบาล	101
70	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG แบ่งตามภาค	103
71	ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS ตั้งแต่ปี 2002-2013 แบ่งตามภาค	104
72	ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG ตั้งแต่ปี 2002-2013 แบ่งตามภาค	105
73	ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 2 waves ตั้งแต่ปี 2002-2013 แบ่งตามภาค	106
74	ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 3 waves ตั้งแต่ปี 2002-2013 แบ่งตามภาค	107
75	ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 4 waves ตั้งแต่ปี 2002-2013 แบ่งตามภาค	108
76	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่อยู่กรุงเทพฯ	109
77	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่อยู่ภาคกลาง	109
78	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่อยู่ภาคเหนือ	110
79	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่อยู่ภาคตะวันออกเฉียงเหนือ	110
80	ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่อยู่ภาคใต้	111
81	จำนวนปีการศึกษาเฉลี่ยรายปีจากข้อมูล LFS และ ORG	112

สารบัญตาราง

ตาราง	หน้า
3.1 จำนวนครัวเรือนตัวอย่างในแผนการเลือกตัวอย่างปี 2545-2546	10
3.2 จำนวนครัวเรือนตัวอย่างของแผนการเลือกตัวอย่างใน samset A	11
3.3 ตัวอย่างตารางข้อมูลตามเงื่อนไขที่ 1	12
3.4 ตัวอย่างตารางข้อมูลตามเงื่อนไขที่ 2	13
3.5 ตัวอย่างตารางข้อมูลตามเงื่อนไขที่ 3	14
3.6 ตัวอย่างตารางข้อมูลตามเงื่อนไขที่ 4	15
3.7 ตัวอย่างตารางข้อมูลตามเงื่อนไขที่ 5	16
3.8 ตัวอย่างตารางข้อมูลที่ต้องมีการสร้าง caseid ใหม่	16
3.9 ตัวอย่างตารางข้อมูลที่มีการสร้าง caseid_new และ caseid_new2	17
4.1 จำนวนตัวอย่างข้าราชการครูและรายบุคคล แบ่งตาม rotation plan	20
4.2 จำนวนตัวอย่างข้าราชการครูที่แบ่งตามเงื่อนไขของอายุ สถานภาพสมรส และจำนวนปีการศึกษา แยกตาม rotation plan	23
4.3 จำนวนตัวอย่างข้าราชการครูที่มีเพศเดียวกัน และอายุห่างกันไม่เกิน 1 ปี สถานภาพสมรสเป็นไปตามเงื่อนไข และจำนวนปีการศึกษาไม่เกิน 1 ปี แบ่งตามวัยแรงงาน	27
4.4 จำนวนตัวอย่างข้าราชการครูที่มีเพศเดียวกัน และอายุห่างกันไม่เกิน 1 ปี สถานภาพสมรสเป็นไปตามเงื่อนไข และจำนวนปีการศึกษาไม่เกิน 1 ปี แบ่งตามจำนวนครั้งที่ถูกสัมภาษณ์	31
4.5 ตัวอย่างตารางข้อมูลที่อาจเกิดข้อผิดพลาดจากการเชื่อมโยงด้วย caseid_new	34

สารบัญตาราง

ตาราง		หน้า
4.6	จำนวนคนที่อยู่ในครัวเรือนเดียวกัน มีเพศและอายุเท่ากัน และถูกสัมภาษณ์รอบเดียวกัน สำหรับกรณีที่ไม่สามารถเชื่อมโยงคนกลุ่มตัวอย่างซ้ำได้จาก caseid และมีการสร้าง caseid_new ตามเกณฑ์ 1 แยกตาม rotation plan	35
4.7	จำนวนคนที่อยู่ในครัวเรือนเดียวกัน มีเพศและอายุเท่ากัน และถูกสัมภาษณ์รอบเดียวกัน สำหรับกรณีที่ไม่สามารถเชื่อมโยงคนกลุ่มตัวอย่างซ้ำได้จาก caseid และมีการสร้าง caseid_new ตามเกณฑ์ 2 แยกตาม rotation plan	38
1	ค่าสถิติพื้นฐานของตัวแปรอายุจากข้อมูล LFS และ ORG	49

นิยามศัพท์

Out Rotation Group หรือ ORG หมายถึง กลุ่มตัวอย่างที่มีการสำรวจซ้ำตามแผนแบบของการสุ่มตัวอย่างจากประชากร เช่น กลุ่มตัวอย่างที่ถูกเลือกมาจากการถูกสัมภาษณ์เดิมในปีแรกร้อยละ 50 เพื่อจะสัมภาษณ์ซ้ำในปีที่ 2 และกลุ่มตัวอย่างใหม่ที่ถูกเลือกอีกร้อยละ 50 เป็นต้น

Cross sectional data หมายถึง ข้อมูลที่เกิดจากการสำรวจ ณ ช่วงเวลาหนึ่งที่กำหนด เช่น ข้อมูลภาวะการทำงานของประชากร ในปี 2545 ณ ไตรมาสที่ 1 เป็นต้น

Panel data หมายถึง ข้อมูลที่เกิดจากการสำรวจในหลาย ๆ ช่วงเวลา โดยข้อมูลที่สำรวจจะเป็นตัวอย่างซ้ำ เช่น ครวเรือนเดิมที่ถูกสัมภาษณ์ซ้ำในปี 2545 ในไตรมาสที่ 1 และ 2 เป็นต้น

Matching ID หมายถึง ตัวเลขที่สร้างขึ้นจากตัวแปรพิกัดที่อยู่อาศัยและเลขสมาชิกครวเรือน เพื่อให้สามารถระบุรหัสของแต่ละบุคคลได้

2-2-2 Pattern หมายถึง รูปแบบของการสำรวจข้อมูลกลุ่มตัวอย่างซ้ำโดยสำรวจกลุ่มตัวอย่างเดิมเว้นไปอีก 2 ช่วงเวลา และจะสำรวจซ้ำอีก 2 ช่วงเวลา เช่น การสำรวจครวเรือนจากข้อมูลภาวะการทำงานของประชากรในปี 2545 ณ ไตรมาสที่ 1 ส่วนไตรมาสที่ 2 และ 3 จะสำรวจครวเรือนใหม่ และจะกลับมาสำรวจครวเรือนเดิมในไตรมาสที่ 4 และไตรมาสที่ 1 ในปี 2546 เป็นต้น

Rotation sampling หรือ rotation plan หมายถึง แผนแบบหรือรูปแบบการสุ่มกลุ่มตัวอย่างซ้ำจากประชากร

สรง. หมายถึง โครงการสำรวจภาวะการทำงานของประชากร

สสช. หมายถึง สำนักงานสถิติแห่งชาติ

Enumeration Area หรือ EA หมายถึง เขตเจนนับหรืออาณาบริเวณทั้งสิ้นของแต่ละจังหวัด จำแนกตามเขตการปกครอง ซึ่งได้จากการนับจุดโครงการสำรวจพื้นฐานของครวเรือน

PSU หมายถึง เขตปฏิบัติการเก็บข้อมูลซึ่งถูกแบ่งออกเป็นชุดจากเขตทั้งหมดทั่วประเทศ

Identification หมายถึง ตัวแปรที่สามารถระบุกลุ่มตัวอย่างรายบุคคลได้

region หมายถึง ตัวแปรที่เป็นรหัสของภาคซึ่งประกอบด้วยเลข 1 หลัก เช่น 1 คือรหัสของกรุงเทพมหานคร และ 2 คือรหัสของภาคกลาง เป็นต้น

province หมายถึง ตัวแปรที่เป็นรหัสจังหวัดตามกรมการปกครองซึ่งประกอบด้วยเลข 2 หลัก เช่น 10 คือรหัสของจังหวัดกรุงเทพมหานคร เป็นต้น

area หมายถึง ตัวแปรประเภทของพื้นที่ที่ทำการสำรวจครัวเรือนซึ่งประกอบด้วยเลข 1 หลัก ได้แก่ 1 คือในเขตเทศบาล และ 2 คือนอกเขตเทศบาล

psunumber หมายถึง ตัวแปรรหัสเขตปฏิบัติการเก็บข้อมูลซึ่งประกอบไปด้วยเลข 4 หลัก เช่น เขตปฏิบัติงานการเก็บข้อมูลที่ 0012 เป็นต้น

samset หมายถึง ตัวแปรรหัสของชุดครัวเรือนตัวอย่างที่ถูกสุ่มตาม rotation plan ซึ่งมีค่าตัวแปรเป็นตัวอักษร ตั้งแต่ค่าที่ A – J โดยในแต่ละ samset จะถูกแบ่งออกเป็น 4 ชุดย่อย ซึ่งถูกเรียกว่า rotation

rotation หมายถึง ตัวแปรรหัสของชุดครัวเรือนตัวอย่างที่สุ่มตาม samset จากแผนแบบ rotation plan โดยใน 1 samset จะมี 4 rotation ซึ่งจะเป็นตัวเลข 1 หลักมีค่าตั้งแต่ 1 – 4 เช่น ครัวเรือนตัวอย่างใน samset A ที่อยู่ใน rotation ที่ 1 เป็นต้น

hhnumber หมายถึง ตัวแปรรหัสรายครัวเรือนกลุ่มตัวอย่างซึ่งประกอบไปด้วยเลข 2 หลัก

memnumber หมายถึง ตัวแปรรหัสรายบุคคลหรือสมาชิกครัวเรือนตัวอย่างซึ่งประกอบไปด้วยเลข 2 หลัก

demographics หมายถึง ตัวแปรที่ระบุภักดิ์ที่ตั้งของครัวเรือนตัวอย่าง ได้แก่ ภาค จังหวัด อำเภอ ตำบล หมู่บ้าน

caseid หมายถึง ตัวแปรรหัสรายบุคคลที่สร้างมาจากตัวแปร demographics ตามด้วยรหัสครัวเรือนและรหัสสมาชิกครัวเรือน

caseid_new หมายถึง caseid ใหม่ที่สร้างขึ้น เนื่องจาก caseid เดิมไม่สามารถระบุรหัสรายบุคคลซึ่งเป็นคนเดียวกันได้ โดยจะมีการจัดเรียงลำดับสมาชิกครัวเรือนใหม่ตามเพศและอายุ และสร้าง caseid ใหม่จากการต่อเลขลำดับสมาชิกครัวเรือนที่จัดเรียงดังกล่าว ซึ่งกำหนดให้อายุของคนที่ต้องใช้ caseid_new เดียวกันจะต้องมีอายุห่างกันไม่เกิน 2 ปี

caseid_new2 หมายถึง caseid ใหม่ที่สร้างขึ้น เนื่องจาก caseid เดิมไม่สามารถระบุรหัสรายบุคคลซึ่งเป็นคนเดียวกันได้ โดยจะมีการจัดเรียงลำดับสมาชิกครัวเรือนใหม่ตามเพศและอายุ และสร้าง caseid ใหม่จากการต่อเลขลำดับสมาชิกครัวเรือนที่จัดเรียงดังกล่าว ซึ่งกำหนดให้อายุของคนที่ต้องใช้ caseid_new2 เดียวกันจะต้องมีอายุห่างกันไม่เกิน 1 ปี

Deterministic Linkage หมายถึง การเชื่อมโยงข้อมูลกลุ่มตัวอย่างรายบุคคลจากการใช้ ID รายบุคคลที่สร้างขึ้นได้

Probabilistic Linkage หมายถึง การเชื่อมโยงข้อมูลกลุ่มตัวอย่างรายบุคคลที่ไม่สามารถใช้ ID รายบุคคลได้โดยตรง จึงต้องใช้การเชื่อมโยงจากลักษณะที่ใกล้เคียงกันของข้อมูลรายบุคคล เช่น เพศ อายุ สถานภาพสมรส ความสัมพันธ์กับหัวหน้าครัวเรือน เป็นต้น

บทที่ 1

ที่มาและความสำคัญ

โครงการวิจัยนี้มีเป้าหมายเพื่อจัดทำฐานข้อมูลรายบุคคลแบบตัวอย่างซ้ำ จากข้อมูลส่วน Out Rotation Group (ORG) ในข้อมูลการสำรวจภาวะการทำงานของประชากร (Labor Force Survey: LFS) ซึ่งจะเป็นประโยชน์ต่อการศึกษาวิจัยในประเด็นสำคัญต่าง ๆ เช่น การย้ายถิ่นของประชากร การเปลี่ยนงานของแรงงาน พลวัตของการจ้างงาน ผลกระทบของการเปลี่ยนแปลงระบบการกำหนดค่าจ้างขั้นต่ำ เป็นต้น เหตุผลที่สำคัญที่ต้องจัดทำเป็นโครงการวิจัยในครั้งนี้ เพราะข้อมูลที่จัดทำขึ้นมีเพียงตัวบ่งชี้ระดับครัวเรือนเท่านั้นที่สามารถเชื่อมโยงระหว่างรอบได้ แต่ตัวบ่งชี้ระดับรายบุคคลภายในครัวเรือนไม่สามารถใช้ได้ทันที นักวิจัยจำเป็นต้องใช้เวลาและความพยายามอย่างมากที่จะเชื่อมโยงบุคคลเดียวกันระหว่างรอบ

ดังนั้น เพื่อส่งเสริมให้เกิดงานวิจัยด้านเศรษฐศาสตร์แรงงานหรือด้านที่เกี่ยวข้องที่ใช้ข้อมูลตัวอย่างซ้ำ (panel data) และลดต้นทุนในการสร้างข้อมูลให้กับนักวิจัย จึงมีความจำเป็นที่จะต้องสร้างข้อมูลตัวอย่างซ้ำขึ้นและเผยแพร่ให้นักวิจัยได้ใช้อย่างสะดวกและทั่วถึง เช่นเดียวกับหน่วยงานที่ส่งเสริมการทำวิจัยในระดับนานาชาติ เช่น NBER (<http://www.nber.org/data/morg.html>) และ CEPR (<http://ceprdata.org/cps-uniform-data-extracts/cps-outgoing-rotation-group/cps-org-programs/>) ต่างก็ได้สร้างข้อมูลตัวอย่างซ้ำแบบนี้จากข้อมูล Current Population Survey: CPS ของประเทศสหรัฐอเมริกา

ข้อมูลแบบตัวอย่างซ้ำที่มีจำนวนกลุ่มตัวอย่างมากพอที่จะเป็นตัวแทนของประชากรได้นั้น มีความสำคัญอย่างมากกับงานวิจัย เพราะช่วยให้เห็นการเปลี่ยนแปลงพฤติกรรมของกลุ่มตัวอย่างและยังสามารถใช้ในการประเมินนโยบายต่าง ๆ ระหว่างก่อนเริ่มนโยบายและหลังเลิกนโยบาย อย่างไรก็ตาม งานวิจัยเกี่ยวกับแรงงานของประเทศไทยได้รับความสนใจและมีประโยชน์ในเชิงนโยบาย แต่งานวิจัยเหล่านั้นยังมีข้อจำกัด ข้อจำกัดที่เป็นประเด็นสำคัญคือ การขาดข้อมูลแบบตัวอย่างซ้ำ ด้านแรงงานของประชากรไทยซึ่งมีความสำคัญต่อการทำความเข้าใจประเด็นที่สำคัญต่าง ๆ เช่น การย้ายถิ่นของประชากร การเปลี่ยนงานของแรงงาน พลวัตของการจ้างงาน ผลกระทบของการเปลี่ยนแปลงระบบการกำหนดค่าจ้างขั้นต่ำ เป็นต้น เราได้จากงานวิจัยในต่างประเทศ เช่น Lemieux (2006), Autor (2011), Acemoglu and Autor (2011) และ Marquis et al. (2014) ส่งผลให้ในช่วงที่ผ่านมา เราไม่สามารถผลิตงานวิจัยด้านเศรษฐศาสตร์แรงงานที่ใช้ข้อมูลแบบตัวอย่างซ้ำในประเทศไทย ในขณะที่สำนักงานสถิติแห่งชาติได้มีการสำรวจภาวะการทำงานของประชากร โดยใช้ระเบียบการเลือกตัวอย่างแบบ Out Rotation Group Sampling นับตั้งแต่ปี พ.ศ. 2545-2556 กำหนดให้มี

ครัวเรือนตัวอย่างซ้ำระหว่างไตรมาสเดียวกันของปีที่ติดกันถึงร้อยละ 50 ของข้อมูลการสำรวจภาวะการทำงานของประชากร ซึ่งข้อมูลจาก Out Rotation Group (ORG) มีจำนวนมากพอที่จะเป็นตัวแทนของประชากรได้ แต่ที่ผ่านมายังไม่มีงานวิจัยที่ใช้ข้อมูล ORG ที่มาจากข้อมูลการสำรวจภาวะการทำงานของประชากร ซึ่งน่าจะมีความเสี่ยงจากความยุ่งยากในการจัดทำข้อมูลตัวอย่างซ้ำระดับรายบุคคล เนื่องจากข้อมูลที่จัดทำขึ้นมีเพียงตัวบ่งชี้ระดับครัวเรือนเท่านั้นที่สามารถเชื่อมโยงระหว่างรอบได้ ตัวบ่งชี้ระดับรายบุคคลยังไม่สามารถใช้ได้ในทันที นักวิจัยต้องใช้เวลาและความพยายามอย่างมากในการเชื่อมโยงข้อมูลบุคคลเดียวกันระหว่างรอบ โดยข้อมูล ORG ที่เชื่อมโยงกลุ่มตัวอย่างซ้ำ (Merged Outgoing Rotation Group: MORG) สร้างจากการนำข้อมูลของกลุ่มตัวอย่างที่เป็นคน ๆ เดียวกันในชุดข้อมูลการสำรวจ ORG ที่เป็นข้อมูลภาคตัดขวาง (cross sectional data) ณ เวลาต่างกัน มาเชื่อมโยงเข้าด้วยกัน ทำให้ได้ข้อมูลที่มีคุณสมบัติของข้อมูลแบบตัวอย่างซ้ำที่บ่งบอกถึงข้อมูลที่เกี่ยวข้องกับกลุ่มตัวอย่างอันเป็นตัวแทนของประชากรในเวลาใด ๆ และการเปลี่ยนแปลงของกลุ่มตัวอย่างกลุ่มเดิมนั้นตามเวลา ซึ่งเอื้ออำนวยให้สามารถวิเคราะห์การเปลี่ยนแปลงที่เกิดขึ้น โดยใช้ปัจจัยแวดล้อมต่าง ๆ ที่เป็นตัวแปรในชุดการสำรวจ มาเป็นตัวแปรอธิบาย รวมถึงการประยุกต์ใช้ข้อมูลเพื่อตอบคำถามทางการวิจัย ที่ต้องอาศัยคุณสมบัติการเป็นตัวอย่างซ้ำเพื่อทดสอบสมมติฐาน ด้วยเหตุนี้ข้อมูล MORG จึงเป็นข้อมูลที่มีคุณค่าและถูกนำไปใช้ประโยชน์อย่างแพร่หลาย เพื่อส่งเสริมให้เกิดงานวิจัยด้านเศรษฐศาสตร์แรงงานและด้านที่เกี่ยวข้องที่ใช้ข้อมูลตัวอย่างซ้ำ และลดต้นทุนในการสร้างข้อมูลให้กับนักวิจัย จึงมีความจำเป็นต้องจัดทำฐานข้อมูลรายบุคคลแบบตัวอย่างซ้ำจากข้อมูลส่วน Out Rotation Group (ORG) ในข้อมูลการสำรวจภาวะการทำงานของประชากร และเผยแพร่ให้นักวิจัยได้ใช้อย่างสะดวกและทั่วถึง เช่นเดียวกับหน่วยงานที่ส่งเสริมการทำวิจัยในระดับนานาชาติ เช่น สำนักงานการวิจัยเศรษฐศาสตร์แห่งชาติ (NBER) ประเทศสหรัฐอเมริกา เป็นผู้จัดทำและเผยแพร่ข้อมูล ORG ที่เชื่อมโยงตัวอย่างซ้ำจากการสำรวจภาวะประชากร (Current Population Survey: CPS) นอกจากนี้ยังมีความพยายามเชื่อมโยงข้อมูลตัวอย่างซ้ำจากข้อมูล CPS เพื่อใช้ประโยชน์จากข้อมูลที่มีอยู่

โครงการนี้จะนำข้อมูลการสำรวจภาวะการทำงานของประชากร (Labor Force Survey: LFS) นับตั้งแต่ปี พ.ศ. 2545 – 2556 มาสร้างข้อมูลตัวอย่างซ้ำด้านแรงงานของประชากรไทยจากข้อมูล ORG ในข้อมูลการสำรวจภาวะการทำงานของประชากร และจะเผยแพร่ข้อมูลส่วนที่เป็น Matching ID ที่สร้างขึ้น พร้อมทั้ง STATA code และคู่มือการสร้างข้อมูลตัวอย่างซ้ำในเว็บไซต์เพื่อให้นักวิจัยไทยได้ใช้อย่างทั่วถึง และเพื่อช่วยให้นักวิจัยที่มีข้อมูลการสำรวจภาวะการทำงานของประชากรสามารถสร้างข้อมูลได้อย่างสะดวก

บทที่ 2

ทบทวนวรรณกรรม

ข้อมูล ORG ที่เชื่อมโยงกลุ่มตัวอย่างซ้ำ (Merged Outgoing Rotation Group: MORG) คือ ชุดข้อมูล ที่สร้างจากการนำข้อมูลของกลุ่มตัวอย่างที่เป็นคน ๆ เดียวกันในชุดข้อมูลการสำรวจ ORG ที่เป็นข้อมูล ภาคตัดขวาง (cross sectional data) ณ เวลาต่างกัน มาเชื่อมโยงเข้าด้วยกัน ทำให้ได้ข้อมูลที่มีคุณสมบัติ ของข้อมูลภาคตัดขวางทางยาว (panel data) คือบ่งบอกถึงข้อมูลที่เกี่ยวข้องกับกลุ่มตัวอย่างอันเป็นตัวแทน ของประชากรใน เวลาใด ๆ และการเปลี่ยนแปลงของกลุ่มตัวอย่างกลุ่มเดิมนั้นตามเวลา ซึ่งเอื้ออำนวยให้ สามารถวิเคราะห์การเปลี่ยนแปลงที่เกิดขึ้น โดยใช้ปัจจัยแวดล้อมต่าง ๆ ที่เป็นตัวแปรในชุดการสำรวจ มาเป็น ตัวแปรอธิบาย รวมถึงการประยุกต์ใช้ข้อมูลเพื่อตอบคำถามทางการวิจัย ที่ต้องอาศัยคุณสมบัติการเป็นตัวอย่าง ซ้ำเพื่อทดสอบสมมติฐาน ด้วยเหตุนี้ข้อมูล MORG จึงเป็นข้อมูลที่มีคุณค่าและถูกนำไปใช้ประโยชน์อย่าง แพร่หลาย

สำหรับในต่างประเทศ อาทิเช่น ในประเทศสหรัฐอเมริกา มีการจัดทำข้อมูล ORG ที่เชื่อมโยงตัวอย่างซ้ำ จากการสำรวจสถานะประชากร (Current Population Survey: CPS) โดยสำนักงานการวิจัยเศรษฐกิจศาสตร์ แห่งชาติ (National Bureau of Economic Research: NBER) และ CEPR เป็นผู้จัดทำและเผยแพร่ โดยใช้ ข้อมูล CPS ที่ดำเนินการสำรวจโดย NBER เป็นฐานข้อมูลที่ใช้สร้างข้อมูล MORG ขึ้นมา สำหรับกระบวนการ เชื่อมโยงข้อมูลนั้น ได้ใช้โปรแกรมสำหรับการจัดการและวิเคราะห์ข้อมูลคือโปรแกรม STATA ในการคัดเลือก และดึงเอาข้อมูล ORG ที่เป็นส่วนหนึ่งของข้อมูล CPS ออกมาเป็นชุดข้อมูลย่อย ซึ่งอ้างอิงตามการสำรวจในแต่ละรอบ จากนั้นจึงทำการรวมข้อมูลย่อยจากการสำรวจแต่ละรอบเข้าด้วยกัน ข้อมูลที่ใช้มีตั้งแต่ปี ค.ศ. 1979 จนถึงปัจจุบัน

สำหรับกระบวนการเชื่อมโยงตัวอย่างนั้น Madrian and Lefgren (2000) อธิบายว่าข้อมูลที่ได้จากการ เชื่อมโยงตัวอย่างซ้ำนี้ มีลักษณะเป็นข้อมูล longitudinal และกระบวนการเชื่อมโยงตัวอย่างจากชุดข้อมูล CPS เข้าด้วยกันนั้นมิได้ทำจากข้อมูลที่ระบุตัวบุคคลโดยตรง (เช่น ชื่อ-นามสกุล) แต่ต้องเชื่อมโยงจาก คุณสมบัติที่บันทึกไว้ของบุคคลนั้น ซึ่งต้องอาศัยเกณฑ์และดุลพินิจของนักวิจัยในการเชื่อมโยง สำหรับ กระบวนการในการเชื่อมโยงนั้น เริ่มจากแบ่งตัวอย่างออกเป็นกลุ่มตามเวลาหรือรอบที่ทำการสำรวจ จากนั้นจึง ทำการเชื่อมโยงกลุ่มตัวอย่างในรอบการสำรวจที่ต่อเนื่องกันเข้าด้วยกัน (เวลา t และ $t+1$) โดยต้องคำนึงถึง แผนการสุ่มตัวอย่างซึ่งกำหนดรูปแบบการเลือกใช้ชุดตัวอย่าง ORG แต่ละชุดมาเป็นกลุ่มตัวอย่างในการสำรวจ แต่ละรอบ ซึ่งชุดข้อมูลที่จำแนกเป็นกลุ่มตัวอย่างนี้ จะต้องเป็นตัวแปรที่เป็นคุณลักษณะที่ใช้เชื่อมโยงกลุ่ม

ตัวอย่างเข้าด้วยกันรวมอยู่ด้วย อาทิเช่น เพศ อายุ เชื้อชาติ เป็นต้น เมื่อได้ชุดข้อมูลตามที่ต้องการแล้ว ให้ทำการสร้างรหัสระบุตัวตนของกลุ่มตัวอย่างเป็นรายบุคคล โดยใช้ตัวแปรที่ระบุตัวตนของบุคคลนั้นได้ เช่น รหัสพื้นที่ รหัสครัวเรือน ขั้นตอนต่อไปทำการเรียงลำดับข้อมูลรายบุคคลในชุดข้อมูลที่มีรอบการสำรวจติดกัน โดยเรียงตามรหัสระบุตัวตนที่กล่าวมาแล้วข้างต้น แล้วจึงเชื่อมโยงบุคคลที่มีรหัสดังกล่าวเหมือนกันเข้าด้วยกัน พร้อมทั้งใช้ข้อมูลคุณลักษณะส่วนบุคคล เช่น เพื่อให้ได้ข้อมูลภาคตัดขวางทางยาว (panel data)

สำหรับการนำข้อมูล MORG ไปใช้ประโยชน์ในทางการวิจัยนั้น มีผู้นำไปประยุกต์ใช้อย่างแพร่หลาย เช่น Lemieux (2006) ใช้ข้อมูล MORG เพื่อประมาณการค่าความคลาดเคลื่อนของค่าจ้างต่อชั่วโมงในชุดข้อมูล CPS เดือนมีนาคม เทียบกับชุดข้อมูล MORG และ CPS เดือนพฤษภาคม โดยได้เปรียบเทียบค่าจ้างของคน ๆ เดียวกันในชุดข้อมูลที่ต่างกันระหว่างข้อมูลค่าจ้างต่อชั่วโมงในชุดข้อมูล CPS เดือนมีนาคม ที่ได้จากการคำนวณและเป็นข้อมูลย้อนหลัง 1 ปี กับค่าจ้างต่อชั่วโมงในชุดข้อมูล ORG และ CPS เดือนพฤษภาคม ที่ได้จากการตอบคำถามโดยตรงและเป็นค่าจ้างในเวลาปัจจุบัน ความคลาดเคลื่อนของค่าจ้างต่อชั่วโมงวัดได้จากผลต่างของค่าความแปรปรวนของค่าจ้างต่อชั่วโมงระหว่างชุดข้อมูลทั้งสองที่กล่าวมาแล้วข้างต้น อีกทั้งงานวิจัยได้ประมาณการ counterfactual ของค่าความคลาดเคลื่อนในกรณีที่สัดส่วนของกลุ่มตัวอย่างไม่เปลี่ยนแปลงตามเวลา เพื่อศึกษาผลของการเปลี่ยนแปลงสัดส่วนของกลุ่มตัวอย่างต่อการเปลี่ยนแปลงของค่าความคลาดเคลื่อนตามเวลา

ผลการเปรียบเทียบแสดงให้เห็นว่าข้อมูลค่าจ้างรายชั่วโมงที่ได้จาก CPS เดือนมีนาคม นั้น มีความคลาดเคลื่อนมากกว่าข้อมูล ORG และข้อมูล CPS เดือนพฤษภาคมพอสมควร อีกทั้งการที่สัดส่วนของกลุ่มตัวอย่างมีแนวโน้มเป็นกลุ่มแรงงานที่ได้รับค่าจ้างรายชั่วโมงมากขึ้น ยิ่งทำให้ค่าความคลาดเคลื่อนของค่าจ้างรายชั่วโมงของชุดข้อมูล CPS เดือนมีนาคม มีแนวโน้มเพิ่มสูงขึ้นตามเวลา ดังนั้นนักวิจัยที่ต้องการใช้ข้อมูลค่าจ้างต่อชั่วโมงควรใช้ข้อมูลนี้จากชุดข้อมูล ORG และ CPS เดือนพฤษภาคม ซึ่งมีความคลาดเคลื่อนน้อยกว่าอันจะทำให้ผลการศึกษาที่มีความถูกต้องมากขึ้น

การเชื่อมโยงชุดข้อมูลโดยใช้ตัวอย่างซ้ำยังสามารถทำให้ได้ชุดข้อมูลที่มีขนาดใหญ่ขึ้น เช่น การเชื่อมโยงชุดข้อมูลโดยใช้ตัวอย่างซ้ำในข้อมูล ORG จะทำให้ได้ชุดข้อมูล MORG ที่มีขนาดใหญ่กว่าข้อมูล CPS เกือบ 3 เท่า ซึ่งเป็นประโยชน์ต่อการวิเคราะห์ทางสถิติ อาทิเช่น ช่วยให้การประมาณการมีความเที่ยงตรงมากขึ้น และเอื้ออำนวยให้สามารถทำการวิเคราะห์ทางสถิติที่ต้องอาศัยจำนวนตัวอย่างขนาดใหญ่ การศึกษาในประเทศสหรัฐอเมริกา ตัวอย่างเช่น Fairlie (2014) ได้ใช้ชุดข้อมูล MORG ตั้งแต่ปี ค.ศ. 1979 – 1998 เพื่อศึกษาแนวโน้มการประกอบอาชีพอิสระของกลุ่มเชื้อชาติต่าง ๆ ซึ่งมีความจำเป็นต้องใช้ชุดข้อมูลที่มีจำนวนตัวอย่างขนาดใหญ่เพื่อวิเคราะห์แนวโน้มและอัตราส่วนของการประกอบอาชีพอิสระในแต่ละกลุ่มเชื้อชาติได้อย่าง

แม่นยำ วิธีการศึกษาได้ใช้ Decomposition method และสมการถดถอยเชิงเส้น เพื่อแยกอิทธิพลของ คุณลักษณะด้านต่าง ๆ ของแรงงานซึ่งเป็นตัวแปรอิสระ อาทิเช่น การศึกษา และเชื้อชาติ ที่มีต่อตัวแปรตามคือ อัตราส่วนของการประกอบอาชีพอิสระ ผลการศึกษาพบว่ากลุ่มเชื้อชาติและผิวดำมีอัตราการประกอบอาชีพ อิสระเพิ่มขึ้นอย่างรวดเร็ว ซึ่งปัจจัยที่ส่งผลคือการเพิ่มขึ้นของกำลังแรงงานของคนกลุ่มนี้ อีกทั้งยังพบว่าการ เพิ่มขึ้นของระดับการศึกษาและการลดลงของอายุของแรงงานบางกลุ่มมีผลทำให้ความแตกต่างของอัตราการ ประกอบอาชีพอิสระของเชื้อชาติต่าง ๆ เพิ่มขึ้น

ข้อมูล MORG ซึ่งเป็นส่วนที่คัดออกมาจากชุดข้อมูล CPS นี่เป็นที่ยอมรับนำไปใช้ในด้านเศรษฐศาสตร์ แรงงานอย่างกว้างขวาง เพราะมีตัวแปรที่เกี่ยวข้องทั้งทางด้านสภาวะการทำงานและคุณลักษณะส่วนบุคคล คร่าว ๆ และปัจจัยที่เกี่ยวข้องอื่น ๆ นอกจากนี้การเชื่อมโยงตัวอย่างซ้ำที่เป็นคน ๆ เดียวกันได้นี้ยัง เอื้ออำนวยให้สามารถสร้างข้อมูลที่มีลักษณะเป็นข้อมูลตัดขวางทางยาว (panel data) ซึ่งสามารถทำการ วิเคราะห์ทางสถิติโดยควบคุมปัจจัยที่เกี่ยวข้อง (fixed effect estimation) ซึ่งมีประโยชน์ในการช่วยให้ สามารถวิเคราะห์ทางสถิติได้อย่างเที่ยงตรงยิ่งขึ้น ตัวอย่างงานวิจัยที่ประยุกต์ใช้คุณสมบัตินี้ เช่น Hoynes et. al. (2012) ได้ศึกษาผลกระทบของการถดถอยทางเศรษฐกิจและวัฏจักรเศรษฐกิจต่อตลาดแรงงานในประเทศ สหรัฐอเมริกา โดยการวิเคราะห์ได้พิจารณาถึงผลกระทบที่มีต่อบุคคลที่มีคุณลักษณะแตกต่างกันในด้านอายุ การศึกษา เพศ และเชื้อชาติ การศึกษาได้ใช้ข้อมูล Merged Outgoing Rotation Group (MORG) ระหว่างปี 2002 -2007 ซึ่งการใช้ข้อมูลที่มีลักษณะเป็น Panel data นี้ได้เอื้ออำนวยให้สามารถวิเคราะห์ข้อมูลโดยใช้ แบบจำลอง state panel model โดยประมาณการสมการถดถอยและมี fixed effect เป็นตัวแปรรัฐ (state) และเวลา (time) โดยกระบวนการ fixed effect โดยใช้ state fixed effect จะทำการลดความแปรปรวนของ ค่าพารามิเตอร์ที่ประมาณค่าได้อันเกิดจากความแตกต่างระหว่างรัฐ และกระบวนการ fixed effect โดยใช้ เวลา จะทำการลดความแปรปรวนของค่าพารามิเตอร์ที่ประมาณค่าได้อันเกิดจากความผันแปรหรือความ แตกต่างของเวลา อาทิเช่น การเปลี่ยนแปลงของอัตราการเข้าสู่ตลาดแรงงานของผู้หญิงระหว่างเวลา และการ ที่วัฏจักรธุรกิจมีผลกระทบต่อรัฐแต่ละรัฐต่างกัน เป็นต้น ทั้งนี้ผลการศึกษาพบว่าผลกระทบของเศรษฐกิจ ถดถอยครั้งใหญ่ในอเมริกานั้นมีความแตกต่างกันในกลุ่มตัวอย่างที่มีคุณลักษณะส่วนบุคคลต่างกัน โดยพบว่า ผลกระทบนั้นสูงสุดในกลุ่มผู้ชาย กลุ่มผิวดำ กลุ่มฮิสแปนิก แรงงานอายุน้อย และการศึกษาต่ำ

ในบางกรณีการสร้างข้อมูล panel data จากการเชื่อมโยงข้อมูลคน ๆ เดียวกันไม่เอื้ออำนวย เช่น กรณีที่ปัญหาทางการวิจัยมีความจำเป็นต้องใช้ข้อมูลที่มีลักษณะเป็น panel data แต่ข้อมูลที่มีเป็นข้อมูล cross-sectional นักวิจัยได้พยายามแก้ปัญหาโดยการสร้างข้อมูลภาคตัดขวางทางยาวเทียม (pseudo-panel) ขึ้นมา โดยข้อมูล pseudo-panel นี้สร้างขึ้นจากการเชื่อมโยงกลุ่มตัวอย่างที่มีคุณลักษณะส่วนบุคคล

เหมือนกันในปีต่างกันเข้าด้วยกัน ทำให้ข้อมูลตัวอย่างซ้ำของคนที่มีลักษณะใกล้เคียงกัน แต่ไม่ใช่คน ๆ เดียวกัน ซึ่งช่วยแก้ข้อจำกัดทางด้านข้อมูลของการศึกษาวิจัยได้ส่วนหนึ่ง การศึกษาวิจัยที่ใช้ข้อมูล pseudo-panel ตัวอย่างเช่น Warunsiri and Mcnown (2010) ได้สร้างข้อมูล pseudo-panel ขึ้นจากข้อมูลภาคตัดขวางคือ ข้อมูลการสำรวจภาวะการทำงานที่ต่อเนื่องกัน ตั้งแต่ปี พ.ศ. 2529 – 2548 เพื่อใช้ประโยชน์จากคุณสมบัติการเป็น panel data ในการประมาณค่าสมการรายได้ Mincer (1974) แบบ fixed effect โดยการประมาณการด้วยวิธีดังกล่าว จะช่วยแก้ไขการเอนเอียง (bias) ของผลตอบแทนของการศึกษาที่ประมาณค่าได้อันเกิดจากปัญหา endogeneity ที่มาจากตัวแปรความสามารถ (ability) ไม่ถูกสำรวจในแบบจำลอง

ข้อมูล panel data ยังถูกใช้ประโยชน์ในการประเมินผลกระทบจากนโยบายค่าแรงขั้นต่ำ ทั้งนี้ การศึกษาผลกระทบของนโยบายค่าแรงขั้นต่ำที่ใช้ข้อมูลรูปแบบอื่นนอกเหนือจาก panel data อาทิเช่น ข้อมูลอนุกรมเวลา (time-series) มักประสบปัญหาหลายประการเช่น พบว่าตัวแปรที่ใช้เป็นตัวแทนของค่าแรงขั้นต่ำ มีการเปลี่ยนแปลงน้อยตลอดช่วงเวลา และการเปลี่ยนแปลงที่เกิดขึ้นมีความสัมพันธ์กับโครงการสวัสดิการสังคมและการฝึกอบรมของรัฐ นอกจากนี้การใช้ข้อมูล cross-sectional data เพื่อศึกษาผลกระทบของค่าแรงขั้นต่ำนั้น อาจมีความคลาดเคลื่อนจากปัจจัยทางเศรษฐกิจของภาคหรือเขตการปกครองนั้น ๆ มีผลทำให้ค่าแรงสูงขึ้นได้ ซึ่งการใช้ข้อมูล panel data จะช่วยให้สามารถควบคุมผลกระทบดังกล่าว รวมถึงปัจจัยอื่น ๆ ที่เกี่ยวข้อง เช่น ปัจจัยด้านภูมิศาสตร์ วัฒนธรรมธุรกิจ กลุ่มประชากรตามรุ่น เป็นต้น (Neumark and Washer, 1992)

บทที่ 3

กระบวนการจัดทำฐานข้อมูลรายบุคคลแบบตัวอย่างซ้ำ จากข้อมูลส่วน Out Rotation Group (ORG)

นักวิจัยได้ศึกษาระเบียบวิธีการเลือกตัวอย่างแบบ Rotation Group ของโครงการสำรวจภาวะการทำงานของประชากรโดยศึกษาจากแผนการเลือกตัวอย่างแบบ 2-2-2 Pattern

สำนักงานสถิติแห่งชาติได้ริเริ่มนำวิธีการเลือกตัวอย่างแบบ Rotation Sampling มาประยุกต์ใช้เพื่อเพิ่มคุณภาพตัวประมาณในโครงการสำรวจภาวะการทำงานของประชากร (สรง.) โดยเริ่มใช้ตั้งแต่ปี พ.ศ. 2545 เป็นต้นมา โดยมีการเก็บข้อมูลทุกเดือนในทุกจังหวัด ทั้งในเขต/นอกเขตเทศบาล มีจำนวนครัวเรือนตัวอย่างประมาณ 84,000 ครัวเรือนต่อไตรมาส เพื่อปรับปรุงคุณภาพของการสำรวจ และเพื่อไม่ให้เป็นการรบกวนครัวเรือนที่จะถูกสัมภาษณ์ซ้ำหลายครั้ง จึงได้มีการพัฒนาการเลือกตัวอย่างแบบ Rotation Sampling เป็น 2 ช่วงที่สำคัญ โดยกำหนดช่วงละ 10 ปี คือ ช่วงเริ่มต้น (ปี พ.ศ. 2545-2554) และช่วงปัจจุบัน (ปี พ.ศ. 2555-2564) รายละเอียดดังนี้

1. ในการเลือกเขตปฏิบัติการเก็บข้อมูล (EA: Enumeration Area) จะมีตัวอย่างรวมทั้งสิ้น 30,000 เขต ซึ่งถูกแบ่งออกเป็น 10-11 ชุด ๆ ละประมาณ 3,000 เขต จากเขตทั้งหมดทั่วประเทศ 128,000 เขต โดยในแต่ละปีของการสำรวจ จะเลือกเขตปฏิบัติการเก็บข้อมูลมา 2 ชุด รวมเป็น 6,000 เขต และในปีที่ติดกันจะมีการใช้ชุดตัวอย่างซ้ำ 1 ชุด คิดเป็น 50% เช่น ในปี พ.ศ. 2545 และ พ.ศ. 2546 จะใช้ชุดเดียวกัน 50% และอีก 50% จะใช้อีกชุดหนึ่ง นอกจากนี้ในแต่ละเขตจะทำการเลือกครัวเรือนตัวอย่าง 4 ชุดย่อย
2. เมื่อสิ้นสุดช่วงแรก (ปี พ.ศ. 2545-2554) สสช. จะทำการเลือกเขตปฏิบัติการเก็บข้อมูลหรือ EA ตัวอย่างทั้งสิ้นของช่วงปัจจุบัน (ปี พ.ศ. 2555-2564) ใหม่ทั้งชุด (11 ชุดย่อย) โดยเริ่มขบวนการทางระเบียบวิธีสถิติใหม่ทั้งหมด ซึ่งทำให้ EA ตัวอย่าง และครัวเรือนตัวอย่างที่ได้เป็นชุดใหม่ ไม่ต่อเนื่อง/ซ้ำกับชุดเดิม (ปี พ.ศ. 2545-2554)
3. จึงทำให้ในช่วงเปลี่ยนผ่านปี 2554 และ 2555 จะไม่มีครัวเรือนตัวอย่างซ้ำ
4. การเลือกตัวอย่างแบบ Rotation Sampling ของ สสช. เป็นแบบ 2-2-2 กล่าวคือ ใช้ครัวเรือนตัวอย่างซ้ำกันใน 2 ไตรมาส และเว้นไปอีก 2 ไตรมาส และซ้ำกันอีก 2 ไตรมาส ซึ่งจะทำให้สามารถติดตามครัวเรือนตัวอย่างซ้ำเพียง 4 ครั้งในรอบ 2 ปีเท่านั้น หลังจากนั้น ครัวเรือนดังกล่าวจะไม่ตกเป็นตัวอย่างอีกเลยในรอบ 10 ปี ตัวอย่างเช่น ครัวเรือนตัวอย่างจะถูกเลือกซ้ำในปี พ.ศ.

2545 ไตรมาสที่ 1 และ 2 และจะเว้นอีก 2 ไตรมาส (ไตรมาสที่ 3 และ 4) หลังจากนั้นครัวเรือนตัวอย่างจะถูกเลือกซ้ำอีกในปี พ.ศ. 2546 ไตรมาสที่ 1 และ 2 ตามลำดับ

โครงการสำรวจภาวะการทำงานของประชากร
แผนภาพแสดงการเลือกตัวอย่างแบบ 4 - Rotation Groups และ 2 - 2 - 2 Pattern

- ชุดใหม่ - เอกสารแนบ 2

ปี	ไตรมาส	Sample and Rotation Group											
		A	B	C	D	E	F	G	H	I	J	K	L
2545	Q ₁	1 2											
	Q ₂	2 3	1										
	Q ₃	3 4	1 2										
	Q ₄	1	4	2 3									
2546	Q ₁	1 2											
	Q ₂	2 3	1										
	Q ₃	3 4	1 2										
	Q ₄	1	4	2 3									
2547	Q ₁	2				3 4							
	Q ₂	3				4							
	Q ₃	4											
	Q ₄												
2548	Q ₁					1							
	Q ₂					2							
	Q ₃					3							
	Q ₄					4							
2549	Q ₁					1							
	Q ₂					2							
	Q ₃					3							
	Q ₄					4							
2550	Q ₁					1							
	Q ₂					2							
	Q ₃					3							
	Q ₄					4							
2551	Q ₁					1							
	Q ₂					2							
	Q ₃					3							
	Q ₄					4							
2552	Q ₁					1							
	Q ₂					2							
	Q ₃					3							
	Q ₄					4							
2553	Q ₁					1							
	Q ₂					2							
	Q ₃					3							
	Q ₄					4							

หมายเหตุ

- ระหว่างไตรมาส (Q) ที่ติดกันจะมีครัวเรือนตัวอย่างซ้ำกัน 50% ยกเว้น Q, 2546 : Q, 2547 } จะมีครัวเรือนตัวอย่างซ้ำกัน 25 %
Q, 2547 : Q, 2547 }
- ไตรมาสเดียวกันในปีที่ติดกันจะมีครัวเรือนตัวอย่างซ้ำกันระหว่าง 0 - 100 %
- ในปี พ.ศ. ที่ลงท้ายด้วยเลขที่ เริ่มตั้งแต่ปี 2547 PSU ในแต่ละ Rotation Group จะเริ่มทยอยถูกทดแทน

รูปที่ 3.1: การเลือกตัวอย่างแบบ 4 Rotation groups และ 2-2-2 pattern ที่ใช้ในปี พ.ศ. 2545 – 2554

เนื่องจากแผนการเลือกตัวอย่างช่วงแรกได้สิ้นสุดลง ในปี พ.ศ. 2554 ดังนั้น จึงได้มีการกำหนดแผนการเลือกตัวอย่างแบบ Rotation ของ สรท. สำหรับใช้ตั้งแต่ปี พ.ศ. 2555-2564 ขึ้นใหม่ โดยแผนการเลือกตัวอย่างรูปแบบใหม่นี้ ได้มีการแบ่ง PSU (EA ตัวอย่าง) ออกเป็น 11 ชุด ในแต่ละ PSU กำหนดให้เลือกครัวเรือนตัวอย่างจำนวน 4 ชุด ซึ่งจะได้ครัวเรือนทั้งสิ้น 44 ชุด โดยในแต่ละปี กำหนดให้ใช้ PSU ในการปฏิบัติงาน จำนวน 2 ชุด

โดยแผนการเลือกตัวอย่างแบบ Rotation ดังกล่าว กำหนดให้มีรูปแบบการซ้ำแบบ 2-2-2 pattern กล่าวคือ

- กำหนดให้ครัวเรือนตัวอย่างถูกเจนนับติดกัน 2 ไตรมาส จากนั้นเว้นไป 2 ไตรมาส และถูกเจนนับซ้ำกันอีก 2 ไตรมาส

- กำหนดให้ PSU ซ้ำกัน 100% และครัวเรือนตัวอย่างซ้ำกัน 50% ระหว่างไตรมาสที่ติดกันภายในปีเดียวกัน
- กำหนด PSU ซ้ำกัน 50% และครัวเรือนตัวอย่างซ้ำกัน 50% ระหว่างไตรมาสเดียวกันของปีที่ติดกัน

โครงการสำรวจภาวะการทำงานของประชากร พ.ศ. 2555-2564

รูปที่ 3-1 : แผนการเลือกตัวอย่าง แบบ 2-2-2 Pattern (มี EA ตัวอย่าง 11 ชุด ในแต่ละ EA มีครัวเรือนตัวอย่าง 4 ชุด) โครงการสำรวจภาวะการทำงานของประชากร พ.ศ. 2555 - 2564

ปี	ไตรมาส	1	2	3	4	5	6	7	8	9	10	11
พ.ศ.	(Q)	A B C D	A B C D	A B C D	A B C D	A B C D	A B C D	A B C D	A B C D	A B C D	A B C D	A B C D
2555	Q ₁	1A 1B 1C 1D	2A 2B 2C 2D									
Q ₂												
Q ₃												
Q ₄												
2556	Q ₁											
Q ₂												
Q ₃												
Q ₄												
2557	Q ₁											
Q ₂												
Q ₃												
Q ₄												
2558	Q ₁											
Q ₂												
Q ₃												
Q ₄												
2559	Q ₁											
Q ₂												
Q ₃												
Q ₄												
2560	Q ₁											
Q ₂												
Q ₃												
Q ₄												
2561	Q ₁											
Q ₂												
Q ₃												
Q ₄												
2562	Q ₁											
Q ₂												
Q ₃												
Q ₄												
2563	Q ₁											
Q ₂												
Q ₃												
Q ₄												
2564	Q ₁											
Q ₂												
Q ₃												
Q ₄												
ตามการนับปี		ค.ศ. - ค.ศ. 54	ค.ศ. - ค.ศ. 55	ค.ศ. - ค.ศ. 56	ค.ศ. - ค.ศ. 57	ค.ศ. - ค.ศ. 58	ค.ศ. - ค.ศ. 59	ค.ศ. - ค.ศ. 60	ค.ศ. - ค.ศ. 61	ค.ศ. - ค.ศ. 62	13 Listing สป. 63	

รูปที่ 3.2: การเลือกตัวอย่างแบบ 4 Rotation groups และ 2-2-2 pattern ที่ใช้ในปี พ.ศ. 2555 – 2564

เมื่อนักวิจัยได้ทำการศึกษาแผนการเลือกแบบ Rotation ข้างต้นแล้วนั้น นักวิจัยได้สร้าง ID ระดับครัวเรือนในแต่ละไตรมาสและแต่ละปี ซึ่งประกอบไปด้วย ตัวแปร region, province, area, psunumber, samset, hhnumber และ ID ระดับบุคคลในแต่ละไตรมาสของแต่ละปี ซึ่งประกอบไปด้วย ตัวแปร region, province, area, psunumber, samset, hhnumber และ memnumber เพื่อค้นหาครัวเรือนตัวอย่างซ้ำในแต่ละ Rotation และเปรียบเทียบกับจำนวนแบบแผนการเลือกตัวอย่างในปี พ.ศ. 2545-2554 และปี พ.ศ. 2555-2556 พบว่า ตัวแปร Identification ของครัวเรือนที่สร้างขึ้นสอดคล้องกับครัวเรือนตัวอย่างตามแบบแผนการเลือกตัวอย่าง 2-2-2 pattern ซึ่งมีครัวเรือนตัวอย่างถูกแฉงนับติดกัน 2 ไตรมาส จากนั้นเว้นไป 2 ไตรมาส และถูกแฉงนับซ้ำกันอีก 2 ไตรมาส จากตัวอย่างตารางแผนการเลือกตัวอย่าง ซึ่งแสดงได้ดังตารางที่

ตารางที่ 3.1: จำนวนครัวเรือนตัวอย่างในแผนการเลือกตัวอย่างปี พ.ศ. 2545-2546

samset	A					B			
year	Round/Rotation	1	2	3	4	1	2	3	4
2545	1	14,929	15,093					14,793	15,007
	2		14,886	14,630		14,726			14,818
	3			14,839	14,601	14,884	14,616		
	4	14,910			14,728		14,750	14,751	
2546	Round/Rotation	1	2	3	4	1	2	3	4
	1	14,863	14,917					14,638	14,852
	2		14,599	14,419		14,455			14,500
	3			14,502	14,406	14,556	14,440		
	4	14,629			14,540		14,563	14,522	

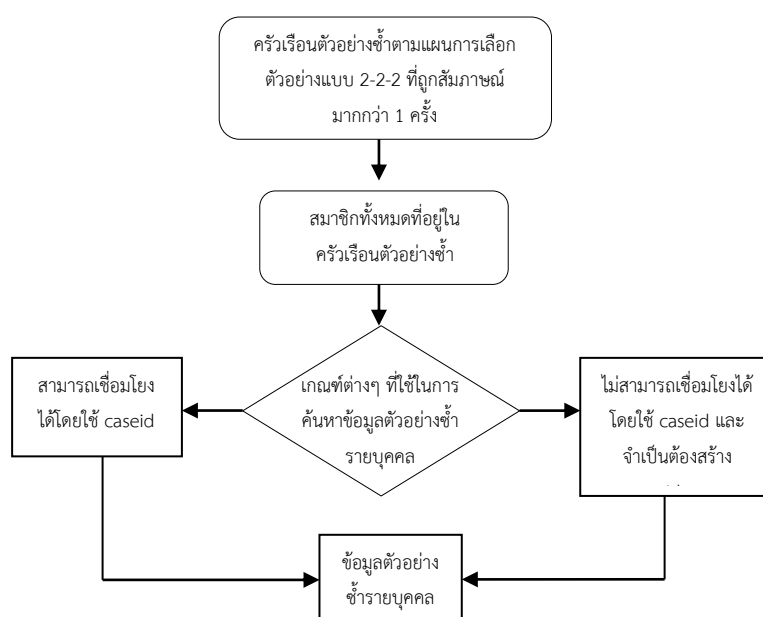
ขั้นตอนการสร้างชุดข้อมูลตามแผนการเลือกตัวอย่างแบบ 2-2-2 pattern

1. จัดการข้อมูลดิบจากการตั้งชื่อตัวแปรและแก้ไขชนิดของตัวแปรให้เหมือนกันทุกไตรมาสและทุกปี (ตั้งแต่ปี พ.ศ. 2545-2556) แล้วเก็บข้อมูลแยกไว้เป็นรายไตรมาสและรายปี
2. เมื่อได้ไฟล์ที่มีการจัดการตัวแปรในขั้นที่ 1 แล้ว หลังจากนั้นเราจะรวมข้อมูลในทุกไตรมาสและทุกปี เพื่อแบ่งตามแผนการเลือกตัวอย่างแบบ 2-2-2 ของ NSO
3. เมื่อได้ข้อมูลในขั้นที่ 2 ตั้งแต่ปี พ.ศ. 2545-2555 แล้วนั้น ขั้นต่อมาเราจะนับจำนวนครัวเรือนตามแผนการเลือกตัวอย่างแบบ 2-2-2 เพื่อสร้างตารางตามแผนแบบดังกล่าว
4. เมื่อได้จำนวนครัวเรือนตัวอย่างตามแผนการเลือกตัวอย่างแล้ว ทำการใส่จำนวนครัวเรือนลงในตารางที่แบ่งตามกลุ่มของ samset/rotation (samset A-J สำหรับข้อมูลในปี พ.ศ. 2545-2554 และ rotation 1-3 สำหรับข้อมูลในปี พ.ศ. 2555-2556)

ตารางที่ 3.2: จำนวนครัวเรือนตัวอย่างของแผนการเลือกตัวอย่างใน samset A

samset	A				
year	Round	Rotation 1	Rotation 2	Rotation 3	Rotation 4
2545	1	14,929	15,093		
	2		14,886	14,630	
	3			14,839	14,601
	4	14,910			14,728
2546	Round	Rotation 1	Rotation 2	Rotation 3	Rotation 4
	1	14,863	14,917		
	2		14,599	14,419	
	3			14,502	14,406
	4	14,629			14,540

5. เมื่อได้จำนวนครัวเรือนตัวอย่างตามแผนแบบแล้วนั้น เราจะทำการค้นหาข้อมูลตัวอย่างซ้ำรายบุคคลในแต่ละ samset และ rotation ตามแผนการเลือกตัวอย่างแบบ 2-2-2 โดยนักวิจัยได้เชื่อมโยงข้อมูลตัวอย่างซ้ำรายบุคคลจาก 2 กรณี คือ กรณีที่ข้อมูลตัวอย่างซ้ำดังกล่าวสามารถใช้ ID ระดับบุคคลได้ (ตั้งชื่อว่า caseid) กับกรณีที่ข้อมูลตัวอย่างซ้ำรายบุคคลที่ไม่สามารถเชื่อมโยงโดยใช้ ID ระดับบุคคลได้ และจำเป็นต้องมีการสร้าง ID ระดับบุคคลขึ้นมาใหม่ (ตั้งชื่อว่า caseid_new) โดยมีขั้นตอนการค้นหาข้อมูลตัวอย่างซ้ำรายบุคคล ตามรูปที่ 3.3 ดังนี้



รูปที่ 3.3: การเชื่อมโยงข้อมูลตัวอย่างซ้ำรายบุคคลจากข้อมูลการสำรวจภาวะการทำงานของประชากร

โดยเกณฑ์ที่นักวิจัยได้กำหนดเพื่อใช้ค้นหาตัวอย่างซ้ำรายบุคคล แบ่งออกเป็น 5 เกณฑ์ ดังนี้

- 1) กลุ่มตัวอย่างซ้ำคนเดียวกันต้องมีเพศเดียวกัน และอายุไม่เกิน 2 ปี นับตั้งแต่การถูกสัมภาษณ์ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่สัมภาษณ์

ตัวอย่างเช่น ตารางที่ 3.3 จะเห็นได้ว่า ในครัวเรือนที่ 1 สมาชิกคนที่ 1 เพศชาย อายุ 45 ปี ในปี พ.ศ. 2545 ไตรมาส 1 และถูกสัมภาษณ์ครั้งที่ 2 คือปี พ.ศ. 2545 ไตรมาส 4 สมาชิกคนที่ 1 เพศชาย มีอายุ 46 ปี และถูกสัมภาษณ์ครั้งที่ 3 และ 4 สมาชิกคนที่ 1 เพศชาย มีอายุเป็น 47 ปี ซึ่งยังคงเป็นไปตามเงื่อนไขที่ 1 ดังนั้น จึงสรุปได้ว่าสมาชิกคนดังกล่าวเป็นคนเดียวกัน

ตารางที่ 3.3: ตัวอย่างตารางข้อมูลตามเงื่อนไขที่ 1

ครัวเรือนที่	ลำดับที่สมาชิก	ปี	ไตรมาส	ครั้งที่สัมภาษณ์	เพศ	อายุ	คนเดียวกัน
1	1	2545	1	1	ชาย	45	ใช่
1	1	2545	4	2	ชาย	46	ใช่
1	1	2546	1	3	ชาย	47	ใช่
1	1	2546	4	4	ชาย	47	ใช่
1	2	2545	1	1	หญิง	35	ใช่
1	2	2545	4	2	หญิง	35	ใช่
1	2	2546	1	3	หญิง	36	ใช่
1	2	2546	4	4	ชาย	33	ไม่ใช่

แต่เมื่อพิจารณาข้อมูลตัวอย่างซ้ำในแต่ละ rotation plan พบว่า ตัวอย่างซ้ำจะถูกสัมภาษณ์ไม่เกิน 1 ปี เช่น ข้อมูลตัวอย่างซ้ำใน samset A จะมีแค่ในปี พ.ศ. 2545 และ พ.ศ. 2546 เท่านั้น ดังนั้น หากใช้เกณฑ์ที่ให้คนเดียวกันมีอายุห่างกันไม่เกิน 2 ปีจากการถูกสัมภาษณ์ครั้งแรกไปจนถึงครั้งสุดท้าย ก็อาจจะไม่ถูกต้องทั้งหมด ดังนั้น นักวิจัยจึงกำหนดเกณฑ์ที่รัดกุมขึ้น ตามเกณฑ์ที่ 2 ดังนี้

- 2) กลุ่มตัวอย่างซ้ำคนเดียวกันต้องมีเพศเดียวกัน และอายุไม่เกิน 1 ปี นับตั้งแต่การถูกสัมภาษณ์ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่สัมภาษณ์

จากครัวเรือนตัวอย่างในตารางที่ 3.4 จะเห็นได้ว่า ครัวเรือนที่ 1 สมาชิกคนที่ 1 เพศชาย ถูกสัมภาษณ์ครั้งที่ 1 มีอายุ 45 ปี และเมื่อถูกสัมภาษณ์ในครั้งที่ 2 ก็ยังเป็นเพศชาย มีอายุ 46 ปี ซึ่งเป็นไปตามเงื่อนไข แต่เมื่อถูกสัมภาษณ์ในครั้งที่ 3 พบว่า สมาชิกคนที่ 1 ยังคงเป็นเพศชาย แต่มีอายุ 47 ปี ซึ่งไม่สอดคล้องกับเงื่อนไขข้างต้น เพราะมีอายุห่างจากการถูกสัมภาษณ์ในครั้งที่ 1 ถึง 2 ปี ดังนั้น หากเราใช้เกณฑ์ที่ 2 ในการ

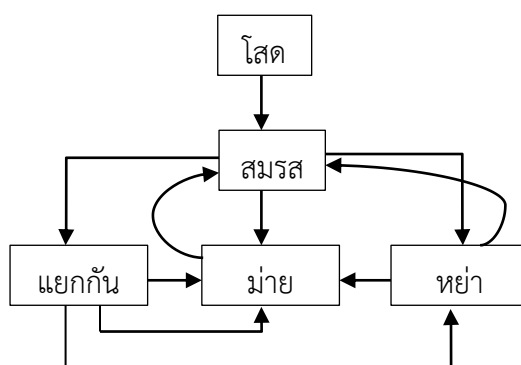
พิจารณา เราจะสรุปว่า สมาชิกคนที่ 1 ที่ถูกสัมภาษณ์ครั้งที่ 3 ไม่ใช่คนเดียวกันกับสมาชิกคนที่ 1 ที่ถูกสัมภาษณ์ครั้งที่ 1 และ 2

ตารางที่ 3.4: ตัวอย่างตารางข้อมูลตามเงื่อนไขที่ 2

ครัวเรือนที่	ลำดับที่สมาชิก	ปี	ไตรมาส	ครั้งที่สัมภาษณ์	เพศ	อายุ	คนเดียวกัน
1	1	2545	1	1	ชาย	45	ใช่
1	1	2545	4	2	ชาย	46	ใช่
1	1	2546	1	3	ชาย	47	ไม่ใช่
1	1	2546	4	4	ชาย	47	ไม่ใช่
1	2	2545	1	1	หญิง	35	ใช่
1	2	2545	4	2	หญิง	35	ใช่
1	2	2546	1	3	หญิง	36	ใช่
1	2	2546	4	4	ชาย	33	ไม่ใช่

นอกจากนี้ นักวิจัยได้ใช้เกณฑ์ที่ 2 เพื่อสร้างเงื่อนไขที่รัดกุมขึ้นโดยเพิ่มตัวแปรของสถานภาพสมรส และกำหนดให้เป็นเกณฑ์ที่ 3 โดยมีเงื่อนไข ดังนี้

- 3) กลุ่มตัวอย่างซ้ำคนเดียวกันต้องมีเพศเดียวกัน และอายุไม่เกิน 1 ปี นับตั้งแต่การถูกสัมภาษณ์ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่สัมภาษณ์ อีกทั้งสถานภาพสมรสต้องเป็นไปตามเงื่อนไข เช่น คนที่โสดต้องไม่เปลี่ยนสถานะเป็นแยกกันอยู่, ม่าย หรือหย่า แต่ควรจะเปลี่ยนเป็นสมรสได้เท่านั้น ตามรูปที่ 3.4 ดังนี้



รูปที่ 3.4: เงื่อนไขของการเปลี่ยนสถานภาพสมรส

เมื่อพิจารณาครัวเรือนตัวอย่างที่ 1 ในตารางที่ 3.5 พบว่า สมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 1 มีอายุ 45 ปี เพศชาย สถานภาพโสด และเมื่อถูกสัมภาษณ์ครั้งที่ 2 มีอายุ 46 ปี เพศชาย แต่มีสถานภาพเป็นม่าย ซึ่งเปลี่ยนจากโสดมาเป็นม่ายก็จะไม่ตรงตามเงื่อนไข เพราะไม่ผ่านการสมรสมาก่อน ดังนั้นสมาชิกคนที่ 1 จึงไม่ใช่คนเดียวกันเมื่อพิจารณาตามเงื่อนไขที่ 3

ตารางที่ 3.5: ตัวอย่างตารางข้อมูลตามเงื่อนไขที่ 3

ครัวเรือนที่	ลำดับที่สมาชิก	ปี	ไตรมาส	ครั้งที่สัมภาษณ์	เพศ	อายุ	สถานภาพ	คนเดียวกัน
1	1	2545	1	1	ชาย	45	โสด	ไม่ใช่
1	1	2545	4	2	ชาย	46	ม่าย	ไม่ใช่
1	1	2546	1	3	ชาย	47	โสด	ไม่ใช่
1	1	2546	4	4	ชาย	47	โสด	ไม่ใช่
1	2	2545	1	1	หญิง	35	สมรส	ใช่
1	2	2545	4	2	หญิง	35	สมรส	ใช่
1	2	2546	1	3	หญิง	36	โสด	ไม่ใช่
1	2	2546	4	4	ชาย	33	สมรส	ไม่ใช่

เช่นเดียวกัน นักวิจัยได้ใช้เกณฑ์ที่ 2 เพื่อสร้างเงื่อนไขที่รัดกุมขึ้นโดยเพิ่มตัวแปรของจำนวนปีการศึกษา และตั้งขึ้นมาเป็นเกณฑ์ที่ 4 โดยมีเงื่อนไข ดังนี้

- 4) กลุ่มตัวอย่างเข้าคนเดียวกันต้องมีเพศเดียวกัน และอายุไม่เกิน 1 ปี นับตั้งแต่การถูกสัมภาษณ์ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่สัมภาษณ์ อีกทั้งจำนวนปีการศึกษาจะต้องมีจำนวนเท่าเดิมหรือเพิ่มขึ้นไม่เกิน 1 ปีการศึกษา นับตั้งแต่การถูกสัมภาษณ์ครั้งแรกจนถึงครั้งสุดท้าย

ตัวอย่างเช่น ครัวเรือนที่ 1 ในตารางที่ 3.6 พบว่า สมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 1 มีอายุ 45 ปี เพศชาย โดยตามเงื่อนไขที่ 4 จะไม่พิจารณาสถานภาพสมรส แต่จะดูจำนวนปีการศึกษาแทน นั่นคือ 12 ปี ซึ่งเมื่อสมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 2 มีอายุ 45 ปี เพศชาย และมีจำนวนปีการศึกษาเท่าเดิมคือ 12 ปี ดังนั้นจึงสรุปว่า ถ้าหากใช้ตามเงื่อนไขที่ 4 สมาชิกคนที่ 1 ที่ถูกสัมภาษณ์ครั้งที่ 1 และ 2 จะเป็นคนเดียวกัน

ตารางที่ 3.6: ตัวอย่างตารางข้อมูลตามเงื่อนไขที่ 4

ครัวเรือนที่	ลำดับที่สมาชิก	ปี	ไตรมาส	ครั้งที่สัมภาษณ์	เพศ	อายุ	สถานภาพ	จำนวนปีการศึกษา	คนเดียวกัน
1	1	2545	1	1	ชาย	45	โสด	12	ใช่
1	1	2545	4	2	ชาย	46	ม่าย	12	ใช่
1	1	2546	1	3	ชาย	47	โสด	12	ไม่ใช่
1	1	2546	4	4	ชาย	47	โสด	12	ไม่ใช่
1	2	2545	1	1	หญิง	35	สมรส	9	ไม่ใช่
1	2	2545	4	2	หญิง	35	สมรส	7	ไม่ใช่
1	2	2546	1	3	หญิง	36	โสด	9	ไม่ใช่
1	2	2546	4	4	ชาย	33	สมรส	9	ไม่ใช่

และเกณฑ์ข้อสุดท้าย นักวิจัยจะใช้เงื่อนไขที่รัดกุมที่สุดจากตัวแปรที่สามารถค้นหาข้อมูลตัวอย่างซ้ำได้ นั่นคือ เพศ, อายุ, สถานภาพสมรส และจำนวนปีการศึกษา โดยมีเงื่อนไข ดังนี้

- 5) กลุ่มตัวอย่างซ้ำคนเดียวกันต้องมีเพศเดียวกัน, อายุไม่เกิน 1 ปี นับตั้งแต่การถูกสัมภาษณ์ครั้งแรก ไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์, อายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่สัมภาษณ์, สถานภาพสมรส และจำนวนปีการศึกษาจะต้องมีจำนวนเท่าเดิมหรือเพิ่มขึ้นไม่เกิน 1 ปีการศึกษา นับตั้งแต่การถูกสัมภาษณ์ครั้งแรกจนถึงครั้งสุดท้าย

จากครัวเรือนตัวอย่างในตารางที่ 3.7 พบว่า ครัวเรือนที่ 1 มีสมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 1 อายุ 45 ปี เพศชาย สถานภาพโสด จำนวนปีการศึกษา 12 ปี และเมื่อสมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 2 มีอายุ 46 ปี เพศชาย สถานภาพม่าย จำนวนปีการศึกษา 12 ปี ซึ่งตามเงื่อนไขที่ 5 แล้วนั้นสมาชิกคนดังกล่าวไม่ใช่คนเดียวกัน เนื่องจากสถานภาพสมรสเปลี่ยนจากโสดไปเป็นม่ายโดยไม่ผ่านการสมรส จึงสรุปได้ว่า สมาชิกคนที่ 1 ไม่ใช่คนเดียวกัน

ตารางที่ 3.7: ตัวอย่างตารางข้อมูลตามเงื่อนไขที่ 5

ครัวเรือนที่	ลำดับที่สมาชิก	ปี	ไตรมาส	ครั้งที่สัมภาษณ์	เพศ	อายุ	สถานภาพ	จำนวนปีการศึกษา	คนเดียวกัน
1	1	2545	1	1	ชาย	45	โสด	12	ไม่ใช่
1	1	2545	4	2	ชาย	46	ม่าย	12	ไม่ใช่
1	1	2546	1	3	ชาย	47	โสด	12	ไม่ใช่
1	1	2546	4	4	ชาย	47	โสด	12	ไม่ใช่
1	2	2545	1	1	หญิง	35	สมรส	9	ไม่ใช่
1	2	2545	4	2	หญิง	35	สมรส	7	ไม่ใช่
1	2	2546	1	3	หญิง	36	โสด	9	ไม่ใช่
1	2	2546	4	4	ชาย	33	สมรส	9	ไม่ใช่

และเนื่องจากนักวิจัยได้ทำการตรวจสอบข้อมูลตัวอย่างซ้ำรายบุคคลโดยใช้ ID ระดับครัวเรือนโดยตั้งชื่อว่า caseid ที่สร้างขึ้นจากตัวแปร demographics และ caseid ใหม่ที่สร้างขึ้นจากข้อมูลที่ไม่สามารถใช้ caseid เดิมได้ โดยตั้งชื่อว่า caseid_new เพราะพบว่า อายุและเพศต่างจาก caseid ที่ถูกสัมภาษณ์ในครั้งแรก ตามตัวอย่างในตารางที่ 3.8

ตารางที่ 3.8: ตัวอย่างตารางข้อมูลที่จำเป็นต้องมีการสร้าง caseid ใหม่

ครัวเรือนที่	ลำดับที่สมาชิก	ปี	ไตรมาส	ครั้งที่สัมภาษณ์	เพศ	อายุ	caseid
2	1	2545	1	1	ชาย	55	ใช้แบบเดิม
2	1	2545	4	2	ชาย	34	ต้องสร้างใหม่
2	1	2546	1	3	ชาย	34	ต้องสร้างใหม่
2	1	2546	4	4	ชาย	35	ต้องสร้างใหม่
2	2	2545	1	1	หญิง	29	ใช้แบบเดิม
2	2	2545	4	2	หญิง	29	ใช้แบบเดิม
2	2	2546	1	3	ชาย	17	ต้องสร้างใหม่
2	2	2546	4	4	ชาย	18	ต้องสร้างใหม่

จากตารางที่ 3.8 ในครัวเรือนตัวอย่างที่ 2 จะเห็นได้ว่า สมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 1 มีอายุ 55 ปี เพศชาย และเมื่อสมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 2 มีอายุ 34 ปี เพศชาย และเมื่อถูกสัมภาษณ์ครั้งที่ 3 และ 4 สมาชิกคนที่ 1 มีอายุ 34 ปี เพศชาย และอายุ 35 ปี เพศชาย ตามลำดับ ซึ่งพบว่า สมาชิกคนที่ 1 ที่ถูกสัมภาษณ์ครั้งแรกมีอายุห่างจากการถูกสัมภาษณ์ครั้งที่ 2-4 มากพอสมควร ทำให้สามารถสรุปได้ว่า สมาชิกคน

ที่ 1 ที่ถูกสัมภาษณ์ครั้งที่ 1 ไม่ใช่คนเดียวกับสมาชิกคนที่ 1 ที่ถูกสัมภาษณ์ครั้งที่ 2-4 เพราะฉะนั้น จึงจำเป็นต้องสร้าง caseid ใหม่ให้กับสมาชิกที่ถูกสัมภาษณ์ครั้งที่ 2-4

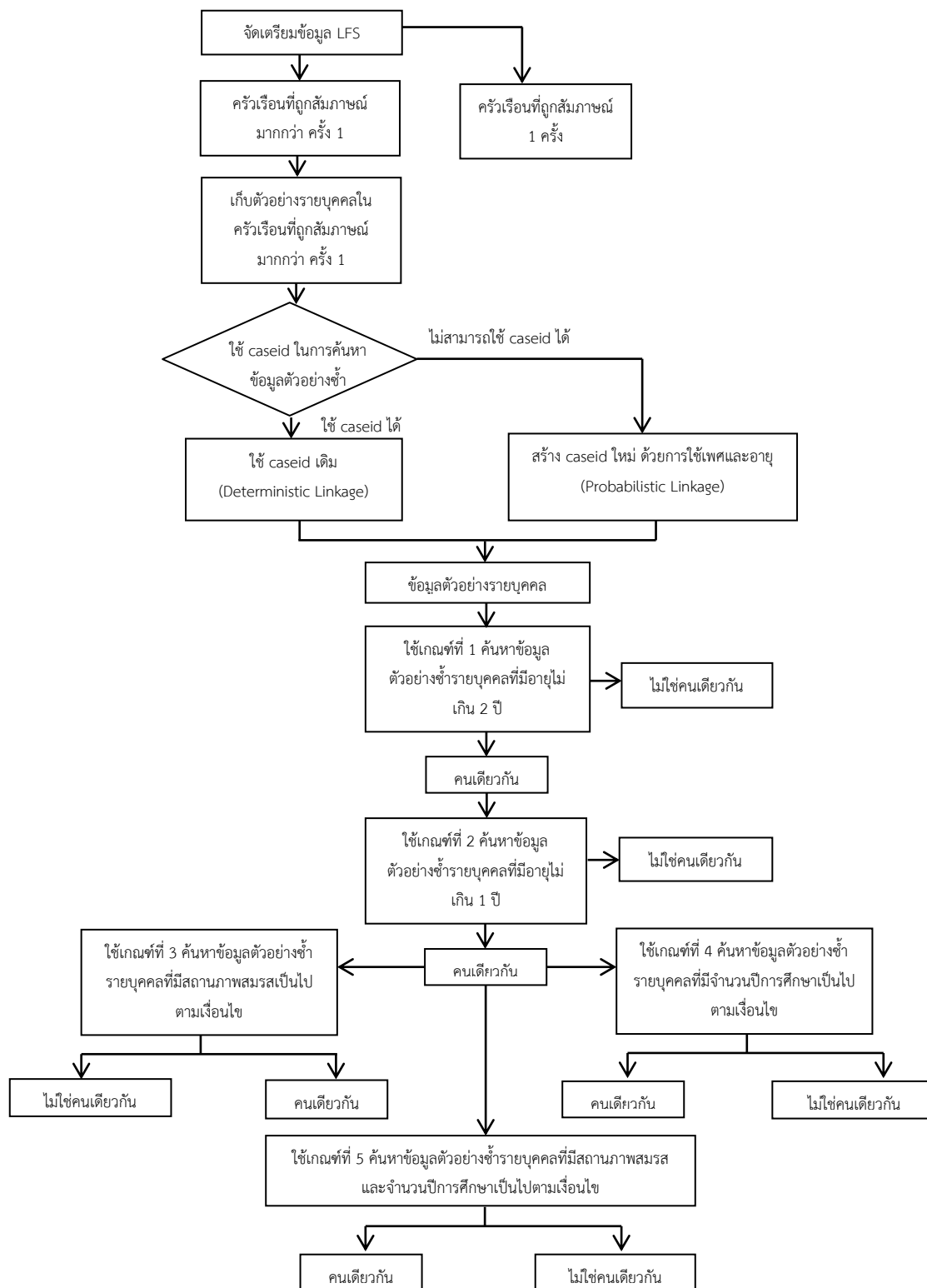
ในการตรวจสอบข้อมูลกลุ่มตัวอย่างซ้ำที่ใช้ caseid ใหม่ ก็ยังคงใช้เกณฑ์ตามเงื่อนไขที่ 1 - 5 เช่นเดียวกับการใช้ caseid แบบเดิม ซึ่งสำหรับการสร้าง caseid ใหม่ นักวิจัยใช้ตัวแปรเพศและอายุในการเชื่อมโยง โดยได้แบ่งการสร้าง caseid ใหม่จาก 2 เงื่อนไข คือ 1) อายุของคนดังกล่าวที่ต้องใช้ caseid_new เดียวกันจะต้องมีอายุห่างกันไม่เกิน 2 ปี (กำหนดเป็น caseid_new) และ 2) อายุของคนดังกล่าวที่ต้องใช้ caseid_new เดียวกันจะต้องมีอายุห่างกันไม่เกิน 1 ปี (กำหนดเป็น caseid_new2) นอกจากนี้ นักวิจัยจะใช้ เลข 2 หลักสุดท้ายของ caseid_new ซึ่งเป็นลำดับที่ต่อจากสมาชิกครัวเรือนสุดท้ายของ caseid เดิม ตามตัวอย่างในตารางที่ 3.9 ดังต่อไปนี้

ตารางที่ 3.9: ตัวอย่างตารางข้อมูลที่มีการสร้าง caseid_new และ caseid_new2

ครัวเรือนที่	ลำดับที่สมาชิก	ปี	ไตรมาส	ครั้งที่สัมภาษณ์	เพศ	อายุ	caseid	caseid_new	caseid_new2
3	1	2545	1	1	ชาย	42	01	01	01
3	1	2545	4	2	ชาย	42	01	01	01
3	1	2546	1	3	ชาย	35	01	03	03
3	1	2546	4	4	ชาย	37	01	03	04
3	2	2545	1	1	หญิง	29	02	02	02
3	2	2545	4	2	หญิง	29	02	02	02
3	2	2546	1	3	ชาย	18	02	04	05
3	2	2546	4	4	ชาย	18	02	04	05

พิจารณาครัวเรือนที่ 3 จากตารางที่ 3.9 พบว่า สมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 1 มีอายุ 42 ปี เพศชาย เมื่อสมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 2 มีอายุ 42 ปี เพศชาย และสมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 3 มีอายุ 35 ปี เพศชาย และสมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 4 มีอายุ 37 ปี เพศชาย ซึ่งจะเห็นได้ว่า สมาชิกคนที่ 1 ที่ถูกสัมภาษณ์ครั้งที่ 3-4 ไม่ใช่คนเดียวกับสมาชิกคนที่ 1 ที่ถูกสัมภาษณ์ครั้งที่ 1-2 ดังนั้น เราจึงต้องสร้าง caseid ใหม่ให้กับสมาชิกคนดังกล่าว และเนื่องจากว่า สมาชิกที่ถูกสัมภาษณ์ครั้งที่ 3 มีอายุ 35 ปี และถูกสัมภาษณ์ครั้งที่ 4 มีอายุ 37 ปี เป็นไปตามเงื่อนไขในการสร้าง caseid_new เพราะมีอายุห่างกันไม่เกิน 2 ปี แต่ถ้าต้องสร้าง caseid_new2 มีเงื่อนไขว่าต้องมีอายุห่างกันไม่เกิน 1 ปี ทำให้สมาชิกคนนี้จะเป็นคนเดียวกันเมื่อใช้ caseid_new แต่จะเป็นคนละคนกันถ้าใช้ caseid_new2

นักวิจัยได้สรุปกระบวนการศึกษาข้อมูลกลุ่มตัวอย่างซ้ำจากข้อมูลภาวะการทำงานของประชากร ตามรูปที่ 3.5 ดังต่อไปนี้



รูปที่ 3.5: กระบวนการจัดทำฐานข้อมูลรายบุคคลแบบตัวอย่างซ้ำจากข้อมูล
ส่วน Out Rotation Group (ORG)

บทที่ 4

ผลการศึกษา

ในบทนี้จะนำเสนอผลลัพธ์จากการทำจัดฐานข้อมูลรายบุคคลแบบตัวอย่างซ้ำจากข้อมูลโครงการสำรวจภาวะการทำงานของประชากร ซึ่งเป็นข้อมูลช่วงปี พ.ศ. 2545 – 2554 และช่วงปี พ.ศ. 2555 – 2556 ด้วยกระบวนการที่ได้อธิบายรายละเอียดไว้ในบทที่ 3 ซึ่งพบว่า มีจำนวนครัวเรือนทั้งหมดโดยเฉลี่ย 18,072 ครัวเรือน แบ่งเป็นจำนวนครัวเรือนที่ถูกสัมภาษณ์เพียง 1 ครั้ง มีเฉลี่ย 892 ครัวเรือน คิดเป็นร้อยละ 5 ของครัวเรือนทั้งหมด และจำนวนครัวเรือนที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง มีเฉลี่ย 17,179 ครัวเรือน คิดเป็นร้อยละ 95 ของครัวเรือนทั้งหมด โดยที่ครัวเรือนที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง มีจำนวนตัวอย่างรายบุคคลทั้งหมดโดยเฉลี่ย 60,870 คน

นอกจากนี้ ข้อมูลรายบุคคลของแต่ละชุดตัวอย่าง ได้ถูกตรวจสอบว่าเป็นตัวอย่างซ้ำหรือไม่ โดยแบ่งเกณฑ์การตรวจสอบข้อมูลโดยใช้ 2 เกณฑ์ คือ 1) กลุ่มตัวอย่างซ้ำที่เป็นคนเดียวกันต้องมีอายุไม่เกิน 2 ปี และ 2) กลุ่มตัวอย่างซ้ำที่เป็นคนเดียวกันต้องมีอายุไม่เกิน 1 ปี ซึ่งพบว่า ตามเกณฑ์ 1 จำนวนตัวอย่างรายบุคคลทั้งหมดที่ถูกสัมภาษณ์ซ้ำอย่างน้อย 1 ครั้ง มีจำนวนเฉลี่ย 53,887 คน คิดเป็นร้อยละ 89 ของจำนวนคนทั้งหมดในแต่ละ rotation plan และตามเกณฑ์ 2 จำนวนตัวอย่างรายบุคคลที่ถูกสัมภาษณ์ซ้ำอย่างน้อย 1 ครั้ง มีจำนวนเฉลี่ย 53,154 คน คิดเป็นร้อยละ 87 ของจำนวนคนทั้งหมดในแต่ละ rotation plan จากนั้นตรวจสอบตัวแปรที่บ่งบอกว่าสมาชิกคนดังกล่าวเป็นคนเดียวกันโดยใช้เพศและอายุ พบว่า ตามเกณฑ์ 1 จำนวนตัวอย่างซ้ำที่เป็นคนเดียวกันมีจำนวนเฉลี่ย 53,435 คน คิดเป็นร้อยละ 88 ของจำนวนคนทั้งหมด และตามเกณฑ์ 2 จำนวนตัวอย่างซ้ำที่เป็นคนเดียวกันมีจำนวนเฉลี่ย 48,262 คน คิดเป็นร้อยละ 79 ของจำนวนคนทั้งหมด ตามลำดับ

อีกทั้ง นักวิจัยได้ตรวจสอบจำนวนตัวอย่างซ้ำรายบุคคลจากตัวแปรสถานภาพสมรส และจำนวนปีการศึกษา โดยพบว่า จำนวนตัวอย่างซ้ำที่มีเพศเดียวกัน อายุห่างกันจากการถูกสัมภาษณ์ครั้งแรกจนถึงครั้งสุดท้ายไม่เกิน 1 ปี สถานภาพสมรสเป็นไปตามเงื่อนไข และจำนวนปีการศึกษาเท่าเดิมหรือเพิ่มขึ้นไม่เกิน 1 ปี มีจำนวนเฉลี่ย 40,457 คน คิดเป็นร้อยละ 66 ของจำนวนคนทั้งหมด ซึ่งสามารถแสดงจำนวนตัวอย่างซ้ำรายบุคคลในแต่ละ rotation plan ได้ดังต่อไปนี้

ตารางที่ 4.1 : จำนวนตัวอย่างซ้ำรายครัวเรือนและรายบุคคล แบ่งตาม rotation plan

ปี	samset	rotation	จำนวน ครัวเรือน ทั้งหมด	จำนวน ครัวเรือนที่ ถูกสัมภาษณ์ มากกว่า 1 ครั้งขึ้นไป	จำนวนคน ทั้งหมด	จำนวนคนที่ ถูกสัมภาษณ์ มากกว่า 1 ครั้งขึ้นไป (ตามเกณฑ์ 1)	จำนวนคนที่ ถูกสัมภาษณ์ มากกว่า 1 ครั้งขึ้นไป (ตามเกณฑ์ 2)
2545-2546	A	1	16,964	15,515	60,730	56,243	55,326
		2	16,894	15,756	60,602	57,472	56,690
		3	16,734	15,547	58,932	55,759	54,932
		4	16,628	15,464	59,152	56,386	55,555
		รวม ร้อยละ	67,220 (100)	62,282 (93)	239,416 (100)	225,860 (94)	222,503 (93)
2545-2546	B	1	16,725	15,552	59,089	56,244	55,382
		2	16,679	15,470	58,702	55,838	55,053
		3	16,826	15,547	59,816	54,876	53,941
		4	16,721	15,689	59,846	56,971	56,165
		รวม ร้อยละ	66,951 (100)	62,258 (93)	237,453 (100)	223,929 (94)	220,541 (93)
2547-2549	C	1	17,101	15,767	60,146	56,339	55,461
		2	17,137	15,845	59,700	55,739	54,973
		3	17,096	15,821	59,384	56,211	55,419
		4	17,286	16,031	60,142	56,423	55,626
		รวม ร้อยละ	68,620 (100)	63,464 (92)	239,372 (100)	224,712 (94)	221,479 (93)
2547-2549	D	1	17,039	15,547	57,516	54,051	53,202
		2	16,981	15,670	58,503	54,875	54,058
		3	16,935	15,607	57,890	54,355	53,537
		4	16,986	15,573	57,866	53,785	52,941
		รวม ร้อยละ	67,941 (100)	62,397 (92)	231,775 (100)	217,066 (94)	213,738 (93)
2549-2551	E	1	18,998	17,935	65,251	62,077	61,247
		2	18,844	17,739	64,355	61,145	60,420
		3	18,491	17,574	62,932	60,219	59,413
		4	18,671	17,686	64,669	60,990	60,054
		รวม ร้อยละ	75,004 (100)	70,934 (95)	257,207 (100)	244,431 (95)	241,134 (94)

ตารางที่ 4.1 (ต่อ)

ปี	samset	rotation	จำนวน ครัวเรือน ทั้งหมด	จำนวน ครัวเรือนที่ ถูกสัมภาษณ์ มากกว่า 1 ครั้งขึ้นไป	จำนวนคน ทั้งหมด	จำนวนคนที่ ถูกสัมภาษณ์ มากกว่า 1 ครั้งขึ้นไป (ตามเกณฑ์ 1)	จำนวนคนที่ ถูกสัมภาษณ์ มากกว่า 1 ครั้งขึ้นไป (ตามเกณฑ์ 2)
2549-2551	F	1	18,421	17,407	63,350	60,307	59,556
		2	18,311	17,370	63,712	60,033	59,219
		3	18,299	17,273	62,652	59,035	58,110
		4	18,150	17,149	62,318	58,901	58,129
		รวม ร้อยละ	73,181 (100)	69,199 (95)	252,032 (100)	238,276 (95)	235,014 (93)
2551-2553	G	1	18,650	17,793	64,732	61,967	61,136
		2	18,741	17,713	64,342	60,717	59,894
		3	18,680	17,822	64,072	61,523	60,689
		4	18,656	17,898	64,809	61,913	61,072
		รวม ร้อยละ	74,727 (100)	67,631 (91)	257,955 (100)	246,120 (95)	242,791 (94)
2551-2553	H	1	18,423	17,550	62,518	59,658	58,905
		2	18,413	17,541	62,325	59,605	58,815
		3	18,463	17,592	63,042	59,703	58,853
		4	18,377	17,386	61,019	56,261	55,447
		รวม ร้อยละ	73,676 (100)	70,069 (95)	248,904 (100)	235,227 (94)	232,020 (93)
2553-2554	I	1	18,001	18,001	64,622	60,044	59,256
		2	18,085	18,011	63,807	56,526	55,670
		3	18,223	17,987	63,009	55,872	55,071
		4	18,255	17,750	61,577	50,724	49,977
		รวม ร้อยละ	72,564 (100)	71,749 (99)	253,015 (100)	223,166 (88)	219,974 (87)
2553-2554	J	1	17,871	17,871	61,544	54,249	53,440
		2	18,099	17,378	59,012	48,058	47,322
		3	18,328	17,066	54,826	37,211	36,912
		4	17,765	17,670	55,164	37,690	37,424
		รวม ร้อยละ	72,063 (100)	68,375 (95)	230,546 (100)	177,208 (77)	175,098 (76)
เฉลี่ยต่อ 1 samset และ rotation ร้อยละ			17,799 (100)	16,839 (95)	61,192 (100)	56,400 (92)	55,607 (91)
รวมทั้งหมด ร้อยละ			711,947 (100)	668,358 (94)	2,447,675 (100)	2,255,995 (92)	2,224,292 (91)

ตารางที่ 4.1 (ต่อ)

ปี	rotation	samset	จำนวน ครัวเรือน ทั้งหมด	จำนวน ครัวเรือนที่ ถูกสัมภาษณ์ มากกว่า 1 ครั้งขึ้นไป	จำนวนคน ทั้งหมด	จำนวนคนที่ ถูกสัมภาษณ์ มากกว่า 1 ครั้งขึ้นไป (ตามเกณฑ์ 1)	จำนวนคนที่ ถูกสัมภาษณ์ มากกว่า 1 ครั้งขึ้นไป (ตามเกณฑ์ 2)
2555	1	A	17,888	17,888	53,129	33,898	33,491
		B	19,844	18,351	56,162	36,569	36,253
		C	19,793	18,299	58,247	39,077	38,783
		D	19,258	18,477	63,054	43,418	43,116
		รวม ร้อยละ	76,783 (100)	73,015 (95)	230,592 (100)	152,962 (66)	151,643 (66)
2555-2556	2	A	17,695	17,695	59,647	52,791	51,725
		B	18,530	17,998	61,652	56,145	55,243
		C	19,204	17,982	64,620	59,050	58,016
		D	19,017	18,298	68,591	65,047	64,152
		รวม ร้อยละ	74,446 (100)	71,973 (97)	254,510 (100)	233,033 (92)	229,136 (90)
2556	3	A	18,262	18,262	53,935	35,670	35,224
		B	19,337	18,828	56,525	40,211	39,995
		C	19,409	18,502	58,571	38,896	38,648
		D	19,549	19,190	63,427	45,345	45,078
		รวม ร้อยละ	76,557 (100)	74,782 (98)	232,458 (100)	160,122 (69)	158,945 (68)
เฉลี่ยต่อ 1 samset และ rotation ร้อยละ			18,072 (100)	17,179 (95)	60,870 (100)	53,887 (89)	53,154 (87)
รวมทั้งหมด ร้อยละ			939,733 (100)	888,128 (95)	3,165,235 (100)	2,802,112 (89)	2,764,016 (87)

จากตารางที่ 4.1 เมื่อพิจารณาข้อมูลตัวอย่างตั้งแต่ปี พ.ศ. 2545-2556 พบว่า จำนวนครัวเรือนตัวอย่าง
 ข้าที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง มีจำนวนเฉลี่ย 17,179 ครัวเรือน คิดเป็นร้อยละ 95 ของจำนวนครัวเรือน
 ทั้งหมดในแต่ละ rotation plan และจำนวนคนที่ถูกสัมภาษณ์มากกว่า 1 ครั้งขึ้นไปตามเกณฑ์ 1 (กลุ่มคน
 ตัวอย่างข้าที่เป็นคนเดียวกันต้องมีอายุไม่เกิน 2 ปี) มีจำนวนเฉลี่ย 53,887 คน คิดเป็นร้อยละ 89 ของจำนวน
 คนทั้งหมด และสำหรับจำนวนคนที่ถูกสัมภาษณ์ 1 ครั้งขึ้นไปตามเกณฑ์ 2 (กลุ่มคนตัวอย่างข้าที่เป็นคน
 เดียวกันต้องมีอายุไม่เกิน 1 ปี) มีจำนวนเฉลี่ย 53,154 คน คิดเป็นร้อยละ 87 ของจำนวนคนทั้งหมด

ทั้งนี้ จะเห็นได้ว่าจำนวนตัวอย่างซ้ำที่อยู่ใน rotation plan สำหรับช่วงปี พ.ศ. 2545-2554 มีร้อยละ โดยเฉลี่ยของจำนวนตัวอย่างซ้ำรายบุคคลสูงกว่าร้อยละโดยเฉลี่ยของจำนวนตัวอย่างซ้ำรายบุคคลที่รวมข้อมูล ตั้งแต่ปี พ.ศ. 2545-2556 เนื่องจาก ใน 2 ปีสุดท้าย คือ ปี พ.ศ. 2555-2556 มีการเปลี่ยนแผนแบบการเลือก ตัวอย่าง ทำให้จำนวนตัวอย่างซ้ำรายบุคคลโดยเฉลี่ยน้อยกว่าปีอื่น ๆ และในบาง rotation plan มีข้อมูล ตัวอย่างที่ถูกสัมภาษณ์ไม่ครบ 4 ครั้งในรอบ 2 ปี (เช่น rotation ที่ 1 samset A ในปี พ.ศ. 2555) ดังนั้น จึง ดึงให้ร้อยละโดยเฉลี่ยของจำนวนตัวอย่างซ้ำรายบุคคล ตั้งแต่ปี พ.ศ. 2545-2556 ต่ำลง จากเหตุผลดังกล่าว นักวิจัยจึงได้แสดงจำนวนตัวอย่างซ้ำรายบุคคลโดยเฉลี่ยต่อ 1 samset และ rotation ออกเป็น 2 ช่วง คือ ปี พ.ศ. 2545-2554 และ ปี พ.ศ. 2545-2556 ตามลำดับ

ตารางที่ 4.2: จำนวนตัวอย่างซ้ำรายบุคคลที่แบ่งตามเงื่อนไขของอายุ สถานภาพสมรส และจำนวนปีการศึกษา แยกตาม rotation plan

ปี	samset	rotation	จำนวนคน ตัวอย่างซ้ำที่ เพศเดียวกัน และอายุห่าง กันไม่เกิน 2 ปี (ตามเกณฑ์ 1)	จำนวนคน ตัวอย่างซ้ำที่ เพศเดียวกัน และอายุห่าง กันไม่เกิน 1 ปี (ตามเกณฑ์ 2)	จำนวนคนตัวอย่าง ซ้ำที่เพศเดียวกัน และอายุห่างกันไม่ เกิน 1 ปี และ สถานภาพสมรส เป็นไปตามเงื่อนไข	จำนวนคนตัวอย่าง ซ้ำที่เพศเดียวกัน และอายุห่างกันไม่ เกิน 1 ปี และ จำนวนปีการศึกษา ไม่เกิน 1 ปี	จำนวนคนตัวอย่างซ้ำที่ เพศเดียวกัน และอายุ ห่างกันไม่เกิน 1 ปี สถานภาพสมรสเป็นไป ตามเงื่อนไข และจำนวน ปีการศึกษาไม่เกิน 1 ปี
2545-2546	A	1	56,243	48,144	47,010	40,273	39,453
		2	57,472	51,059	49,944	42,766	41,937
		3	55,759	50,294	49,161	42,319	41,491
		4	56,386	51,012	49,888	42,933	42,101
		รวม ร้อยละ	225,860 (94)	200,509 (84)	196,003 (82)	168,291 (70)	164,982 (69)
2545-2546	B	1	56,244	50,557	49,469	42,523	41,689
		2	55,838	50,544	49,408	42,465	41,615
		3	54,876	46,798	45,690	39,125	38,320
		4	56,971	50,537	49,407	42,395	41,570
		รวม ร้อยละ	223,929 (94)	198,436 (84)	193,974 (82)	166,508 (70)	163,984 (69)
2547-2549	C	1	56,339	49,673	48,341	41,005	40,018
		2	55,739	49,935	48,561	41,602	40,580
		3	56,211	50,785	49,373	42,466	41,435
		4	56,423	49,697	48,389	41,542	40,574
		รวม ร้อยละ	224,712 (94)	200,090 (84)	194,664 (81)	166,615 (70)	162,607 (68)

ตารางที่ 4.2 (ต่อ)

ปี	samset	rotation	จำนวนคน ตัวอย่างซ้ำที่ เพศเดียวกัน และอายุห่าง กันไม่เกิน 2 ปี (ตามเกณฑ์ 1)	จำนวนคน ตัวอย่างซ้ำที่ เพศเดียวกัน และอายุห่าง กันไม่เกิน 1 ปี (ตามเกณฑ์ 2)	จำนวนคนตัวอย่าง ซ้ำที่เพศเดียวกัน และอายุห่างกันไม่ เกิน 1 ปี และ สถานภาพสมรส เป็นไปตามเงื่อนไข	จำนวนคนตัวอย่าง ซ้ำที่เพศเดียวกัน และอายุห่างกันไม่ เกิน 1 ปี และ จำนวนปีการศึกษา ไม่เกิน 1 ปี	จำนวนคนตัวอย่างซ้ำที่ เพศเดียวกัน และอายุ ห่างกันไม่เกิน 1 ปี สถานภาพสมรสเป็นไป ตามเงื่อนไข และ จำนวนปีการศึกษาไม่ เกิน 1 ปี
2547-2549	D	1	54,051	48,779	47,500	40,726	39,800
		2	54,875	48,272	46,998	40,484	39,542
		3	54,355	48,112	46,851	40,147	39,201
		4	53,785	48,448	47,086	40,688	39,695
		รวม ร้อยละ	217,066 (94)	193,611 (84)	188,435 (81)	162,045 (70)	158,238 (68)
2549-2551	E	1	62,077	55,433	53,870	52,513	51,079
		2	61,145	55,571	54,171	52,667	51,382
		3	60,219	54,968	53,493	52,014	50,667
		4	60,990	54,060	52,676	54,060	52,676
		รวม ร้อยละ	244,431 (95)	220,032 (86)	214,210 (83)	211,254 (82)	205,804 (80)
2549-2551	F	1	60,307	55,036	53,577	52,217	50,880
		2	60,033	53,405	51,974	53,405	51,974
		3	59,035	52,491	51,097	43,488	42,410
		4	58,901	52,847	51,459	44,072	43,023
		รวม ร้อยละ	238,276 (95)	213,779 (85)	208,107 (83)	193,182 (77)	188,287 (75)
2551-2553	G	1	61,967	55,208	53,637	45,923	44,767
		2	60,717	55,059	53,571	46,310	45,233
		3	61,523	56,047	54,476	46,941	45,780
		4	61,913	54,779	53,273	46,011	44,891
		รวม ร้อยละ	246,120 (95)	221,093 (86)	214,957 (83)	185,185 (72)	180,671 (70)
2551-2553	H	1	59,658	54,339	52,854	45,665	44,581
		2	59,605	52,844	51,431	44,534	43,473
		3	59,703	53,251	51,737	44,301	43,143
		4	56,261	51,038	49,596	43,533	42,415
		รวม ร้อยละ	235,227 (95)	211,472 (85)	205,618 (83)	178,033 (72)	173,612 (70)

ตารางที่ 4.2 (ต่อ)

ปี	rotation	samset	จำนวนคน ตัวอย่างซ้ำที่ เพศเดียวกัน และอายุห่าง กันไม่เกิน 2 ปี (ตามเกณฑ์ 1)	จำนวนคน ตัวอย่างซ้ำที่ เพศเดียวกัน และอายุห่าง กันไม่เกิน 1 ปี (ตามเกณฑ์ 2)	จำนวนคนตัวอย่าง ซ้ำที่เพศเดียวกัน และอายุห่างกันไม่ เกิน 1 ปี และ สถานภาพสมรส เป็นไปตามเงื่อนไข	จำนวนคนตัวอย่าง ซ้ำที่เพศเดียวกัน และอายุห่างกันไม่ เกิน 1 ปี และ จำนวนปีการศึกษา ไม่เกิน 1 ปี	จำนวนคนตัวอย่างซ้ำ ที่เพศเดียวกัน และ อายุห่างกันไม่เกิน 1 ปี สถานภาพสมรส เป็นไปตามเงื่อนไข และจำนวนปี การศึกษาไม่เกิน 1 ปี
2553-2554	I	1	59,433	53,250	51,635	44,266	43,098
		2	55,615	49,944	48,401	42,159	41,024
		3	54,853	49,845	48,268	41,674	40,513
		4	50,079	45,929	44,582	39,532	38,510
		รวม ร้อยละ	219,980 (87)	198,968 (79)	192,886 (76)	167,631 (66)	163,145 (64)
2553-2554	J	1	53,154	48,218	46,755	40,463	39,390
		2	47,342	43,329	42,064	37,295	36,329
		3	36,320	34,273	33,657	30,467	29,985
		4	36,353	34,689	34,033	31,262	30,741
		รวม ร้อยละ	173,169 (75)	160,509 (70)	156,509 (68)	139,487 (61)	136,445 (59)
เฉลี่ยต่อ 1 samset และ rotation ร้อยละ			56,219 (92)	50,462 (82)	49,134 (80)	43,456 (71)	42,425 (69)
รวมทั้งหมด ร้อยละ			2,248,770 (92)	2,018,499 (82)	1,965,363 (80)	1,738,231 (71)	1,696,985 (69)

ตารางที่ 4.2 (ต่อ)

ปี	samset	rotation	จำนวนคน ตัวอย่างซ้ำที่ เพศเดียวกัน และอายุห่าง กันไม่เกิน 2 ปี (ตามเกณฑ์ 1)	จำนวนคน ตัวอย่างซ้ำที่ เพศเดียวกัน และอายุห่าง กันไม่เกิน 1 ปี (ตามเกณฑ์ 2)	จำนวนคนตัวอย่าง ซ้ำที่เพศเดียวกัน และอายุห่างกันไม่ เกิน 1 ปี และ สถานภาพสมรส เป็นไปตามเงื่อนไข	จำนวนคนตัวอย่าง ซ้ำที่เพศเดียวกัน และอายุห่างกันไม่ เกิน 1 ปี และ จำนวนปีการศึกษา ไม่เกิน 1 ปี	จำนวนคนตัวอย่างซ้ำ ที่เพศเดียวกัน และ อายุห่างกันไม่เกิน 1 ปี สถานภาพสมรส เป็นไปตามเงื่อนไข และจำนวนปี การศึกษาไม่เกิน 1 ปี
2555	1	A	32,780	29,901	29,316	26,208	25,749
		B	35,440	33,533	32,944	29,881	29,397
		C	37,774	35,991	35,411	32,391	31,931
		D	42,085	40,364	39,794	36,653	36,193
		รวม ร้อยละ	148,079 (64)	139,789 (61)	137,465 (60)	125,133 (54)	123,270 (53)
2555-2556	2	A	51,223	44,291	42,922	35,311	34,356
		B	54,565	49,096	47,668	40,201	39,184
		C	57,389	52,110	50,599	41,139	40,105
		D	63,173	58,360	56,731	45,348	44,253
		รวม ร้อยละ	226,350 (89)	203,857 (80)	197,920 (78)	161,999 (64)	157,898 (62)
2556	3	A	34,682	31,788	31,192	27,009	26,551
		B	39,039	37,500	36,974	34,065	33,645
		C	37,778	36,013	35,323	30,179	29,667
		D	43,927	42,196	41,523	36,240	35,750
		รวม ร้อยละ	155,426 (67)	147,497 (63)	145,012 (62)	127,493 (55)	125,613 (54)
เฉลี่ยต่อ 1 samset และ rotation ร้อยละ			53,435 (88)	48,262 (79)	47,034 (77)	41,401 (68)	40,457 (66)
รวมทั้งหมด ร้อยละ			2,778,625 (88)	2,509,642 (79)	2,445,760 (77)	2,152,856 (68)	2,062,196 (65)

จากตารางที่ 4.2 เมื่อพิจารณาข้อมูลตัวอย่างตั้งแต่ปี พ.ศ. 2545-2556 พบว่า จำนวนตัวอย่างซ้ำรายบุคคลที่มีเพศเดียวกัน และอายุห่างกันไม่เกิน 2 ปี มีจำนวนเฉลี่ย 53,435 คน คิดเป็นร้อยละ 88 ของจำนวนคนทั้งหมด และจำนวนตัวอย่างซ้ำรายบุคคลที่มีเพศเดียวกัน และอายุห่างกันไม่เกิน 1 ปี มีจำนวนเฉลี่ย 48,262 คน คิดเป็นร้อยละ 79 ของจำนวนคนทั้งหมด นอกจากนี้ หากเพิ่มเงื่อนไขของสถานภาพสมรสกับตัวอย่างซ้ำรายบุคคลที่มีเพศเดียวกัน และอายุห่างกันไม่เกิน 1 ปี พบว่า มีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ย 47,034 คน คิดเป็นร้อยละ 77 ของจำนวนคนทั้งหมด และเช่นเดียวกัน ถ้าพิจารณาจำนวนปีการศึกษากับตัวอย่างซ้ำรายบุคคลที่มีเพศเดียวกัน และอายุห่างกันไม่เกิน 1 ปี พบว่า มีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ย

41,401 คน คิดเป็นร้อยละ 68 ของจำนวนคนทั้งหมด อีกทั้งเมื่อพิจารณาทั้งสถานภาพสมรสและจำนวนปีการศึกษากับตัวอย่างข้าราชการบุคคลที่มีเพศเดียวกัน และอายุห่างกันไม่เกิน 1 ปี พบว่า มีจำนวนตัวอย่างข้าราชการบุคคลเฉลี่ย 40,457 คนคิดเป็นร้อยละ 66 ของจำนวนคนทั้งหมด นั่นคือ เมื่อนักวิจัยเพิ่มเงื่อนไขของเกณฑ์การค้นหาข้อมูลตัวอย่างข้าราชการบุคคลให้รัดกุมมากขึ้น ร้อยละของจำนวนตัวอย่างข้าราชการบุคคลก็จะยิ่งลดลงแต่ก็ยังคงมีร้อยละของจำนวนตัวอย่างข้าราชการบุคคลที่มากกว่าร้อยละ 50 ของจำนวนคนทั้งหมด

ตารางที่ 4.3: จำนวนตัวอย่างข้าราชการบุคคลที่มีเพศเดียวกัน และอายุห่างกันไม่เกิน 1 ปี สถานภาพสมรสเป็นไปตามเงื่อนไข และจำนวนปีการศึกษาไม่เกิน 1 ปี แบ่งตามวัยแรงงาน

ปี	samset	rotation	จำนวนคนตัวอย่างซ้ำที่เพศเดียวกัน และอายุห่างกันไม่เกิน 1 ปี สถานภาพสมรสเป็นไปตามเงื่อนไข และจำนวนปีการศึกษาไม่เกิน 1 ปี	จำนวนคนตัวอย่างซ้ำที่อยู่ในวัยแรงงาน (15-59 ปี)	จำนวนคนตัวอย่างซ้ำที่ไม่อยู่ในวัยแรงงาน (น้อยกว่า 15 ปี มากกว่า 59 ปี)
2545-2546	A	1	39,453	22,213	17,240
		2	41,937	23,444	18,493
		3	41,491	23,410	18,081
		4	42,101	23,744	18,357
		รวม ร้อยละ	164,982 (100)	92,811 (56)	72,171 (44)
2545-2546	B	1	41,689	23,484	18,205
		2	41,615	23,233	18,382
		3	38,320	21,631	16,689
		4	41,570	23,259	18,311
		รวม ร้อยละ	163,984 (100)	91,607 (56)	71,587 (44)
2547-2549	C	1	40,018	22,084	17,934
		2	40,580	22,394	18,186
		3	41,435	23,087	18,348
		4	40,574	22,346	18,228
		รวม ร้อยละ	162,607 (100)	89,911 (55)	72,696 (45)
2547-2549	D	1	39,800	22,297	17,503
		2	39,542	21,835	17,707
		3	39,201	21,862	17,339
		4	39,695	22,106	17,589
		รวม ร้อยละ	158,238 (100)	88,100 (56)	70,138 (44)

ตารางที่ 4.3 (ต่อ)

ปี	samset	rotation	จำนวนคนตัวอย่างซ้ำที่เพศเดียวกัน และอายุห่างกันไม่เกิน 1 ปี สถานภาพสมรสเป็นไปตามเงื่อนไข และจำนวนปีการศึกษาไม่เกิน 1 ปี	จำนวนคนตัวอย่างซ้ำที่อยู่ในวัยแรงงาน (15-59 ปี)	จำนวนคนตัวอย่างซ้ำที่ไม่ อยู่ในวัยแรงงาน (น้อยกว่า 15 ปี มากกว่า 59 ปี)
2549-2551	E	1	51,079	29,838	21,241
		2	51,382	30,098	21,284
		3	50,667	29,540	21,127
		4	52,676	31,211	21,465
		รวม ร้อยละ	205,804 (100)	120,687 (59)	85,117 (41)
2549-2551	F	1	50,880	29,863	21,017
		2	51,974	30,697	21,277
		3	42,410	23,479	18,931
		4	43,023	23,668	19,355
		รวม ร้อยละ	188,287 (100)	107,707 (57)	80,580 (43)
2551-2553	G	1	44,767	24,833	19,934
		2	45,233	25,246	19,987
		3	45,780	25,582	20,198
		4	44,891	24,487	20,404
		รวม ร้อยละ	180,671 (100)	100,148 (55)	80,523 (45)
2553-2554	J	1	39,390	21,825	17,565
		2	36,329	20,314	16,015
		3	29,985	16,482	13,503
		4	30,741	17,370	13,371
		รวม ร้อยละ	136,445 (100)	75,991 (56)	60,454 (44)
2551-2553	H	1	44,581	24,914	19,667
		2	43,473	24,132	19,341
		3	43,143	23,810	19,333
		4	42,415	23,831	18,584
		รวม ร้อยละ	173,612 (100)	96,687 (56)	76,925 (44)

ตารางที่ 4.3 (ต่อ)

ปี	rotation	samset	จำนวนคนตัวอย่างซ้ำที่เพศเดียวกัน และอายุห่างกันไม่เกิน 1 ปี สถานภาพสมรสเป็นไปตามเงื่อนไข และจำนวนปีการศึกษาไม่เกิน 1 ปี	จำนวนคนตัวอย่างซ้ำที่อยู่ในวัยแรงงาน (15-59 ปี)	จำนวนคนตัวอย่างซ้ำที่ไม่อยู่ในวัยแรงงาน (น้อยกว่า 15 ปี มากกว่า 59 ปี)
2551-2553	H	1	44,581	24,914	19,667
		2	43,473	24,132	19,341
		3	43,143	23,810	19,333
		4	42,415	23,831	18,584
		รวมร้อยละ	173,612 (100)	96,687 (56)	76,925 (44)
2553-2554	I	1	43,098	23,798	19,300
		2	41,024	22,631	18,393
		3	40,513	22,522	17,991
		4	38,510	21,517	16,993
		รวมร้อยละ	163,145 (100)	90,468 (55)	72,677 (45)
เฉลี่ยต่อ 1 samset และ rotation ร้อยละ			42,425 (100)	23,853 (56)	18,572 (44)
รวมทั้งหมด ร้อยละ			1,696,985 (100)	954,117 (56)	742,868 (44)

ตารางที่ 4.3 (ต่อ)

ปี	rotation	samset	จำนวนคนตัวอย่างซ้ำที่เพศเดียวกัน และอายุห่างกันไม่เกิน 1 ปี สถานภาพ สมรสเป็นไปตามเงื่อนไข และจำนวน ปีการศึกษาไม่เกิน 1 ปี	จำนวนคนตัวอย่างซ้ำที่อยู่ในวัย แรงงาน (15-59 ปี)	จำนวนคนตัวอย่างซ้ำที่ไม่ อยู่ในวัยแรงงาน (น้อยกว่า 15 ปี มากกว่า 59 ปี)
2555	1	A	25,749	14,259	11,490
		B	29,397	16,458	12,939
		C	31,931	17,968	13,963
		D	36,193	20,104	16,089
		รวมร้อยละ	123,270 (100)	68,789 (56)	54,481 (44)
2555-2556	2	A	34,356	19,103	15,253
		B	39,184	21,453	17,731
		C	40,105	21,951	18,154
		D	44,253	23,572	20,681
		รวมร้อยละ	157,898 (100)	86,079 (55)	71,819 (45)
2556	3	A	26,551	14,906	11,645
		B	33,645	19,041	14,604
		C	29,667	16,530	13,137
		D	35,750	19,882	15,868
		รวมร้อยละ	125,613 (100)	70,359 (56)	55,254 (44)
เฉลี่ยต่อ 1 samset และ rotation ร้อยละ			40,457 (100)	22,680 (56)	17,777 (44)
รวมทั้งหมด ร้อยละ			2,103,766 (100)	1,179,344 (56)	924,422 (44)

จากตาราง 4.3 เมื่อพิจารณาข้อมูลตัวอย่างตั้งแต่ปี พ.ศ. 2545-2556 พบว่า จำนวนตัวอย่างซ้ำรายบุคคลที่อยู่ในวัยแรงงานซึ่งมีอายุในช่วง 15-59 ปี มีจำนวนเฉลี่ย 22,680 คน คิดเป็นร้อยละ 56 ของจำนวนตัวอย่างซ้ำรายบุคคลที่มีเพศเดียวกัน และอายุห่างกันไม่เกิน 1 ปี สถานภาพสมรสเป็นไปตามเงื่อนไข และจำนวนปีการศึกษาไม่เกิน 1 ปี ในขณะที่จำนวนตัวอย่างซ้ำรายบุคคลที่ไม่อยู่ในวัยแรงงานมีจำนวนเฉลี่ย 17,777 คน คิดเป็นร้อยละ 44 ของจำนวนตัวอย่างซ้ำรายบุคคลที่มีเพศเดียวกัน และอายุห่างกันไม่เกิน 1 ปี สถานภาพสมรสเป็นไปตามเงื่อนไข และจำนวนปีการศึกษาไม่เกิน 1 ปี

ตารางที่ 4.4: จำนวนตัวอย่างซ้ำรายบุคคลที่มีเพศเดียวกัน และอายุห่างกันไม่เกิน 1 ปี สถานภาพสมรสเป็นไปตามเงื่อนไข และจำนวนปีการศึกษาไม่เกิน 1 ปี แบ่งตามจำนวนครั้งที่ถูกสัมภาษณ์

ปี	samset	rotation	จำนวนคนตัวอย่างซ้ำที่มีเพศเดียวกัน และอายุห่างกันไม่เกิน 1 ปี สถานภาพสมรสเป็นไปตามเงื่อนไข และจำนวนปีการศึกษาไม่เกิน 1 ปี	ถูกสัมภาษณ์ 2 ครั้ง	ถูกสัมภาษณ์ 3 ครั้ง	ถูกสัมภาษณ์ 4 ครั้ง
2545-2546	A	1	39,453	15,296	12,856	11,301
		2	41,937	16,832	11,904	13,201
		3	41,491	17,459	11,844	12,188
		4	42,101	18,000	11,835	12,266
		รวมร้อยละ	164,982 (100)	67,587 (41)	48,439 (29)	48,956 (30)
2545-2546	B	1	41,689	17,410	11,787	12,492
		2	41,615	17,449	11,641	12,525
		3	38,320	15,161	12,386	10,773
		4	41,570	16,742	11,981	12,847
		รวมร้อยละ	163,984 (100)	66,762 (40)	47,795 (30)	48,637 (30)
2547-2549	C	1	40,018	17,284	11,835	10,899
		2	40,580	16,709	11,910	11,961
		3	41,435	16,672	12,140	12,623
		4	40,574	16,376	11,928	12,270
		รวมร้อยละ	162,607 (100)	67,041 (41)	47,813 (30)	47,753 (29)
2547-2549	D	1	39,800	16,101	11,702	11,997
		2	39,542	15,880	11,655	12,007
		3	39,201	15,790	11,602	11,809
		4	39,695	15,939	12,019	11,737
		รวมร้อยละ	158,238 (100)	63,710 (40)	46,978 (30)	47,550 (30)
2549-2551	E	1	51,079	19,789	15,537	15,753
		2	51,382	20,185	15,234	15,963
		3	50,667	18,906	15,062	16,699
		4	52,676	20,132	15,217	17,327
		รวมร้อยละ	205,804 (100)	79,012 (38)	61,050 (30)	65,742 (32)

ตารางที่ 4.4 (ต่อ)

ปี	samset	rotation	จำนวนคนตัวอย่างซ้ำที่ เพศเดียวกัน และอายุห่าง กันไม่เกิน 1 ปี สถานภาพ สมรสเป็นไปตามเงื่อนไข และจำนวนปีการศึกษาไม่ เกิน 1 ปี	ถูกสัมภาษณ์ 2 ครั้ง	ถูกสัมภาษณ์ 3 ครั้ง	ถูกสัมภาษณ์ 4 ครั้ง
2549-2551	F	1	50,880	19,281	14,966	16,633
		2	51,974	19,471	15,301	17,202
		3	42,410	17,016	12,755	12,639
		4	43,023	17,288	12,601	13,134
		รวม ร้อยละ	188,287 (100)	73,056 (38)	55,623 (30)	59,608 (32)
2551-2553	G	1	44,767	17,170	12,895	14,702
		2	45,233	17,522	13,611	14,100
		3	45,780	17,246	13,087	15,447
		4	44,891	16,546	12,599	15,746
		รวม ร้อยละ	180,671 (100)	68,484 (39)	52,192 (29)	59,995 (32)
2551-2553	H	1	44,581	16,791	12,539	15,251
		2	43,473	16,232	12,264	14,977
		3	43,143	16,370	12,377	14,396
		4	42,415	16,941	15,902	9,572
		รวม ร้อยละ	173,612 (100)	66,334 (38)	53,082 (29)	54,196 (33)
2553-2554	I	1	43,098	15,540	12,487	15,071
		2	41,024	16,569	14,496	9,959
		3	40,513	16,597	14,784	9,132
		4	38,510	18,479	20,031	-
		รวม ร้อยละ	163,145 (100)	67,185 (38)	61,798 (31)	34,162 (31)
เฉลี่ยต่อ 1 samset และ rotation ร้อยละ			42,425 (100)	17,847 (42)	12,685 (30)	11,893 (28)
รวมทั้งหมด ร้อยละ			1,696,985 (100)	713,870 (42)	507,395 (30)	475,720 (28)

ตารางที่ 4.4 (ต่อ)

ปี	rotation	samset	จำนวนคนตัวอย่างซ้ำที่ เพศเดียวกัน และอายุห่าง กันไม่เกิน 1 ปี สถานภาพ สมรสเป็นไปตามเงื่อนไข และจำนวนปีการศึกษาไม่ เกิน 1 ปี	ถูกสัมภาษณ์ 2 ครั้ง	ถูกสัมภาษณ์ 3 ครั้ง	ถูกสัมภาษณ์ 4 ครั้ง
2553-2554	J	1	39,390	16,384	13,885	9,121
		2	36,329	17,589	18,740	-
		3	29,985	29,985	-	-
		4	30,741	30,741	-	-
		รวม ร้อยละ	136,445 (100)	94,699 (41)	32,625 (38)	9,121 (21)
2555	1	A	25,749	25,749	-	-
		B	29,397	29,397	-	-
		C	31,931	31,931	-	-
		D	36,193	36,193	-	-
		รวม ร้อยละ	123,270 (100)	123,270 (100)	-	-
2555-2556	2	A	34,356	14,221	9,911	10,224
		B	39,184	16,615	10,715	11,854
		C	40,105	17,451	11,540	11,114
		D	44,253	20,819	10,846	12,588
		รวม ร้อยละ	157,898 (100)	69,106 (44)	43,012 (27)	45,780 (29)
2556	3	A	26,551	26,551	-	-
		B	33,645	33,645	-	-
		C	29,667	29,667	-	-
		D	35,750	35,750	-	-
		รวม ร้อยละ	125,613 (100)	125,613 (100)	-	-
เฉลี่ยต่อ 1 samset และ rotation ร้อยละ			40,457 (100)	19,843 (49)	10,585 (26)	10,029 (25)
รวมทั้งหมด ร้อยละ			2,103,766 (100)	1,031,859 (49)	550,407 (26)	521,500 (25)

จากตาราง 4.4 เมื่อพิจารณาข้อมูลตัวอย่างตั้งแต่ปี พ.ศ. 2545-2556 พบว่า จำนวนตัวอย่างซ้ำรายบุคคลที่ถูกสัมภาษณ์ 2 ครั้ง มีจำนวนเฉลี่ย 19,843 คน คิดเป็นร้อยละ 49 ของจำนวนตัวอย่างซ้ำรายบุคคลที่มีเพศเดียวกัน และอายุห่างกันไม่เกิน 1 ปี สถานภาพสมรสเป็นไปตามเงื่อนไข และจำนวนปี

การศึกษาไม่เกิน 1 ปี ในขณะที่จำนวนคนตัวอย่างซ้ำที่ถูกสัมภาษณ์ 3 ครั้ง มีจำนวนเฉลี่ย 10,585 คน คิดเป็นร้อยละ 26 ของจำนวนตัวอย่างซ้ำรายบุคคลที่มีเพศเดียวกัน และอายุห่างกันไม่เกิน 1 ปี สถานภาพสมรสเป็นไปตามเงื่อนไข และจำนวนปีการศึกษาไม่เกิน 1 ปี และจำนวนตัวอย่างซ้ำรายบุคคลที่ถูกสัมภาษณ์ครบ 4 ครั้ง มีจำนวนเฉลี่ย 10,029 คน คิดเป็นร้อยละ 25 ของจำนวนตัวอย่างซ้ำรายบุคคลที่มีเพศเดียวกัน และอายุห่างกันไม่เกิน 1 ปี สถานภาพสมรสเป็นไปตามเงื่อนไข และจำนวนปีการศึกษาไม่เกิน 1 ปี

ในกรณีที่นักวิจัยไม่สามารถเชื่อมโยงตัวอย่างซ้ำรายบุคคลได้จาก caseid นักวิจัยจะทำการเชื่อมโยงจากเพศและอายุที่ใกล้เคียงกันในรอบการถูกสัมภาษณ์ครั้งถัดไปแทน ซึ่งจะมีกรณีที่ครัวเรือนตัวอย่างซ้ำ มีสมาชิกครัวเรือนที่มีเพศเดียวกัน และอายุเท่ากันในการถูกสัมภาษณ์รอบเดียวกัน (เช่น ฝาแฝด) ทำให้ในรอบถัดไป นักวิจัยอาจเชื่อมโยงคนใดคนหนึ่งสลับกันได้ และเกิดเป็นความผิดพลาดจากการสร้าง caseid ใหม่ขึ้น ตัวอย่างเช่น

ตารางที่ 4.5: ตัวอย่างตารางข้อมูลที่อาจเกิดข้อผิดพลาดจากการเชื่อมโยงด้วย caseid_new

ครัวเรือนที่	ลำดับที่สมาชิก	ปี	ไตรมาส	ครั้งที่สัมภาษณ์	เพศ	อายุ	caseid	caseid_new
4	1	2545	4	1	หญิง	38	01	01
4	2	2545	4	1	ชาย	10	02	02
4	3	2545	4	1	ชาย	10	03	04
4	1	2546	1	2	ชาย	39	01	01
4	2	2546	1	2	ชาย	11	02	02
4	3	2546	1	2	ชาย	21	03	05

จากตารางที่ 4.5 จะเห็นได้ว่า สมาชิกคนที่ 2 และ 3 มีเพศและอายุเท่ากันในการสัมภาษณ์ ครั้งที่ 1 ดังนั้น ในการสัมภาษณ์ครั้งที่ 2 จึงเป็นไปได้ว่าสมาชิกเพศชายที่มีอายุ 11 ปีอาจเป็นสมาชิกคนที่ 2 หรือคนที่ 3 คนใดคนหนึ่ง โดยนักวิจัยตั้งสมมติฐานว่าให้สมาชิกลำดับที่มาก่อนเชื่อมโยงกับสมาชิกในรอบถัดไป นั่นคือให้สมาชิกคนที่ 2 เป็นสมาชิกเพศชายที่มีอายุ 11 ปี ในการสัมภาษณ์ครั้งที่ 2

ในส่วนถัดไป นักวิจัยได้แสดงจำนวนคนกลุ่มตัวอย่างซ้ำที่อาจมีโอกาสเชื่อมโยงผิดพลาดจากการสร้าง caseid_new ตามกรณีข้างต้น ตามตาราง 4.6 - 4.7 ดังนี้

ตารางที่ 4.6: จำนวนคนที่อยู่ในครัวเรือนเดียวกัน มีเพศและอายุเท่ากัน และถูกสัมภาษณ์รอบเดียวกัน สำหรับกรณีที่ไม่สามารถเชื่อมโยงตัวอย่างซ้ำรายบุคคลได้จาก caseid และมีการสร้าง caseid_new ตามเกณฑ์ 1 แยกตาม rotation plan

ปี	samset	rotation	จำนวนคนที่ไม่สามารถเชื่อมโยงได้ด้วย caseid และมีการสร้าง caseid_new ด้วยเกณฑ์ 1 (คนเดียวกันต้องมีอายุห่างกันไม่เกิน 2 ปี)	สัมภาษณ์ ในครั้งที่ 1	สัมภาษณ์ ในครั้งที่ 2	สัมภาษณ์ ในครั้งที่ 3	สัมภาษณ์ ในครั้งที่ 4
2545-2546	A	1	49,340	197	243	260	278
		2	46,215	186	144	235	202
		3	45,494	146	260	229	271
		4	46,444	190	202	213	247
		รวม ร้อยละ	187,493 (100)	719 (0.38)	849 (0.45)	937 (0.50)	998 (0.53)
2545-2546	B	1	45,834	148	278	232	253
		2	45,670	183	218	205	219
		3	49,376	183	250	223	290
		4	45,630	149	142	221	208
		รวม ร้อยละ	186,510 (100)	663 (0.36)	888 (0.48)	881 (0.47)	970 (0.52)
2547-2549	C	1	49,350	216	221	284	186
		2	47,356	172	240	221	218
		3	47,130	206	214	237	193
		4	46,660	223	237	247	218
		รวม ร้อยละ	190,496 (100)	817 (0.43)	912 (0.48)	989 (0.52)	815 (0.43)
2547-2549	D	1	44,659	224	274	227	205
		2	45,296	190	265	233	187
		3	44,997	178	166	191	154
		4	45,682	160	233	174	183
		รวม ร้อยละ	180,634 (100)	752 (0.42)	938 (0.52)	825 (0.46)	729 (0.40)
2549-2551	E	1	46,236	167	197	203	209
		2	46,727	155	178	170	177
		3	44,116	149	184	159	182
		4	47,879	158	240	227	196
		รวม ร้อยละ	184,958 (100)	629 (0.34)	799 (0.43)	759 (0.41)	764 (0.41)

ตารางที่ 4.6 (ต่อ)

ปี	samset	rotation	จำนวนคนที่ไม่สามารถ เชื่อมโยงได้ด้วย caseid และ มีการสร้าง caseid_new ด้วย เกณฑ์ 1 (คนเดียวกันต้องมี อายุห่างกันไม่เกิน 2 ปี)	สัมภาษณ์ ในครั้งที่ 1	สัมภาษณ์ ในครั้งที่ 2	สัมภาษณ์ ในครั้งที่ 3	สัมภาษณ์ ในครั้งที่ 4
2549-2551	F	1	45,409	188	205	188	204
		2	47,270	144	184	218	235
		3	47,881	158	170	174	165
		4	45,784	132	172	189	250
		รวม ร้อยละ	186,344 (100)	622 (0.33)	731 (0.39)	769 (0.41)	854 (0.46)
2551-2553	G	1	46,713	169	158	159	180
		2	48,068	144	191	189	204
		3	45,795	214	194	188	156
		4	45,015	148	166	180	155
		รวม ร้อยละ	185,591 (100)	675 (0.36)	709 (0.38)	716 (0.39)	695 (0.37)
2551-2553	H	1	44,445	175	186	209	156
		2	43,596	114	173	183	173
		3	45,472	191	178	207	145
		4	40,906	100	178	161	96
		รวม ร้อยละ	174,419 (100)	580 (0.33)	135 (0.48)	760 (0.44)	570 (0.33)
2553-2554	I	1	47,232	165	188	235	186
		2	46,439	161	132	200	211
		3	46,189	152	216	227	227
		4	39,021	224	194	175	-
		รวม ร้อยละ	178,881 (100)	702 (0.39)	730 (0.41)	837 (0.47)	624 (0.35)
2553-2554	J	1	45,961	121	197	110	204
		2	37,593	188	184	181	-
		3	32,463	262	175	-	-
		4	32,795	217	237	-	-
		รวม ร้อยละ	148,812 (100)	788 (0.53)	793 (0.53)	401 (0.27)	204 (0.14)
เฉลี่ยต่อ 1 samset และ rotation ร้อยละ			45,103 (100)	174 (0.39)	202 (0.45)	204 (0.45)	201 (0.44)
รวมทั้งหมด ร้อยละ			1,804,138 (100)	6,947 (0.39)	7,484 (0.41)	7,874 (0.44)	7,223 (0.40)

ตารางที่ 4.6 (ต่อ)

ปี	rotation	samset	จำนวนคนที่ไม่สามารถ เชื่อมโยงได้ด้วย caseid และมี การสร้าง caseid_new ด้วย เกณฑ์ 1 (คนเดียวกันต้องมี อายุห่างกันไม่เกิน 2 ปี)	สัมภาษณ์ ในครั้งที่ 1	สัมภาษณ์ ในครั้งที่ 2	สัมภาษณ์ ในครั้งที่ 3	สัมภาษณ์ ในครั้งที่ 4
2555	1	A	34,558	225	142	-	-
		B	36,160	204	165	-	-
		C	35,917	245	177	-	-
		D	36,987	317	271	-	-
		รวม ร้อยละ	143,622 (100)	991 (0.69)	755 (0.53)	-	-
2555-2556	2	A	54,164	169	118	105	123
		B	54,601	175	147	111	116
		C	58,030	165	152	140	196
		D	59,911	206	216	213	172
		รวม ร้อยละ	226,706 (100)	715 (0.32)	633 (0.44)	569 (0.40)	607 (0.42)
2556	3	A	33,198	169	107	-	-
		B	30,255	130	138	-	-
		C	36,758	233	171	-	-
		D	35,083	238	216	-	-
		รวม ร้อยละ	135,294 (100)	770 (0.57)	632 (0.47)	-	-
เฉลี่ยต่อ 1 samset และ rotation ร้อยละ			44,418 (100)	181 (0.41)	194 (0.45)	198 (0.45)	196 (0.44)
รวมทั้งหมด ร้อยละ			2,309,760 (100)	9,423 (0.41)	9,504 (0.41)	8,443 (0.37)	7,830 (0.34)

จากตารางที่ 4.6 แสดงจำนวนคนที่ไม่สามารถเชื่อมโยงได้ด้วย caseid และมีการสร้าง caseid ใหม่ด้วยเกณฑ์ 1 (คนเดียวกันต้องมีอายุห่างกันไม่เกิน 2 ปี) ซึ่งเมื่อพิจารณาข้อมูลตัวอย่างตั้งแต่ปี พ.ศ. 2545-2556 มีจำนวนเฉลี่ย 44,418 คน เมื่อพิจารณาจำนวนคนที่อยู่ในครัวเรือนเดียวกัน มีเพศและอายุเท่ากัน และถูกสัมภาษณ์รอบที่ 1 มีจำนวนเฉลี่ย 181 คน คิดเป็นร้อยละ 0.41 นั้นหมายความว่า การใช้ caseid ใหม่ที่สร้างขึ้นตามเกณฑ์ 1 มีโอกาสเกิดข้อผิดพลาดประมาณร้อยละ 0.4

ตารางที่ 4.7: จำนวนคนที่อยู่ในครัวเรือนเดียวกัน มีเพศและอายุเท่ากัน และถูกสัมภาษณ์รอบเดียวกัน สำหรับกรณีที่ไม่สามารถเชื่อมโยงตัวอย่างซ้ำรายบุคคลได้จาก caseid และมีการสร้าง caseid_new ตามเกณฑ์ 2 แยกตาม rotation plan

ปี	samset	rotation	จำนวนคนที่ไม่สามารถเชื่อมโยงได้ด้วย caseid และมีการสร้าง caseid_new2 ด้วยเกณฑ์ 2 (คนเดียวกันต้องมีอายุห่างกันไม่เกิน 1 ปี)	สัมภาษณ์ ในครั้งที่ 1	สัมภาษณ์ ในครั้งที่ 2	สัมภาษณ์ ในครั้งที่ 3	สัมภาษณ์ ในครั้งที่ 4
2545-2546	A	1	53,044	211	256	253	304
		2	49,070	204	153	229	201
		3	48,156	144	255	225	299
		4	49,017	184	222	207	248
		รวม ร้อยละ	199,287 (100)	743 (0.37)	886 (0.44)	914 (0.46)	1,052 (0.53)
2545-2546	B	1	48,571	159	278	230	255
		2	48,067	187	221	215	237
		3	53,018	198	265	220	302
		4	48,463	155	145	231	214
		รวม ร้อยละ	198,119 (100)	699 (0.35)	909 (0.46)	896 (0.50)	1,008 (0.43)
2547-2549	C	1	52,455	201	231	283	196
		2	50,267	188	241	213	237
		3	49,723	220	211	252	218
		4	49,605	214	240	257	221
		รวม ร้อยละ	202,050 (100)	823 (0.41)	923 (0.47)	1,005 (0.50)	872 (0.43)
2547-2549	D	1	47,184	223	272	239	205
		2	48,145	195	268	249	203
		3	47,888	187	173	201	162
		4	48,224	166	234	182	194
		รวม ร้อยละ	191,441 (100)	771 (0.48)	947 (0.49)	871 (0.45)	764 (0.40)
2549-2551	E	1	49,075	185	194	218	220
		2	49,292	162	177	186	181
		3	46,480	158	184	163	186
		4	50,760	165	260	240	200
		รวม ร้อยละ	195,607 (100)	670 (0.34)	815 (0.42)	807 (0.41)	787 (0.40)

ตารางที่ 4.7 (ต่อ)

ปี	samset	rotation	จำนวนคนที่ไม่สามารถ เชื่อมโยงได้ด้วย caseid และมี การสร้าง caseid_new2 ด้วย เกณฑ์ 2 (คนเดียวกันต้องมี อายุห่างกันไม่เกิน 1 ปี)	สัมภาษณ์ ในครั้งที่ 1	สัมภาษณ์ ในครั้งที่ 2	สัมภาษณ์ ในครั้งที่ 3	สัมภาษณ์ ในครั้งที่ 4
2549-2551	F	1	47,924	182	218	192	199
		2	50,026	151	213	217	237
		3	50,851	157	169	195	184
		4	48,446	137	177	187	267
		รวม ร้อยละ	197,247 (100)	627 (0.32)	777 (0.39)	791 (0.40)	887 (0.45)
2551-2553	G	1	49,702	168	183	168	189
		2	50,758	151	192	193	216
		3	48,292	207	198	206	174
		4	47,820	157	170	187	168
		รวม ร้อยละ	196,572 (100)	683 (0.35)	743 (0.38)	754 (0.38)	747 (0.38)
2551-2553	H	1	46,827	177	179	221	168
		2	46,252	121	174	179	185
		3	48,306	194	175	219	147
		4	42,990	113	134	162	98
		รวม ร้อยละ	184,375 (100)	605 (0.33)	662 (0.36)	781 (0.42)	598 (0.32)
2553-2554	I	1	49,895	167	193	236	203
		2	48,779	166	128	209	234
		3	48,193	154	231	236	231
		4	40,207	213	195	192	-
		รวม ร้อยละ	187,074 (100)	700 (0.37)	747 (0.40)	873 (0.47)	668 (0.36)
2553-2554	J	1	47,955	131	199	229	205
		2	38,724	189	184	185	-
		3	32,762	263	181	-	-
		4	33,061	218	241	-	-
		รวม ร้อยละ	152,502 (100)	801 (0.53)	805 (0.53)	414 (0.27)	205 (0.13)
เฉลี่ยต่อ 1 samset และ rotation ร้อยละ			47,607 (100)	178 (0.37)	205 (0.43)	203 (0.43)	190 (0.40)
รวมทั้งหมด ร้อยละ			1,904,274 (100)	7,122 (0.37)	8,214 (0.43)	8,106 (0.43)	7,588 (0.40)

ตารางที่ 4.7 (ต่อ)

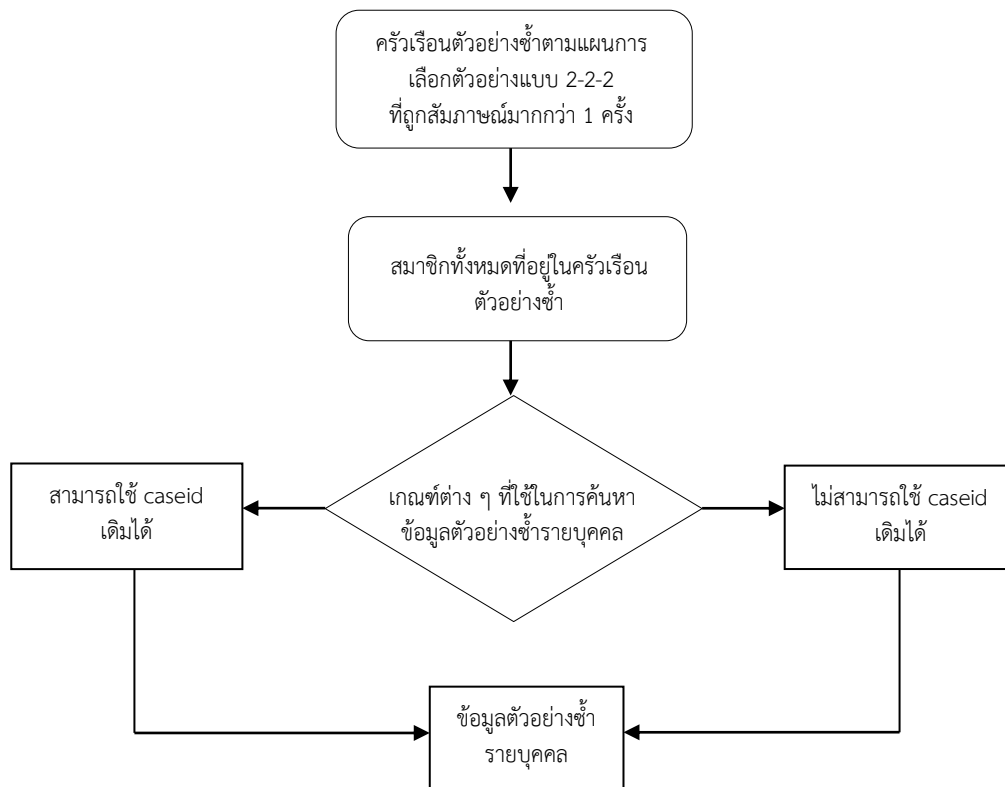
ปี	rotation	samset	จำนวนคนที่ไม่สามารถเชื่อมโยงได้ด้วย caseid และมีการสร้าง caseid_new2 ด้วยเกณฑ์ 2 (คนเดียวกันต้องมีอายุห่างกันไม่เกิน 1 ปี)	สัมภาษณ์ ในครั้งที่ 1	สัมภาษณ์ ในครั้งที่ 2	สัมภาษณ์ ในครั้งที่ 3	สัมภาษณ์ ในครั้งที่ 4
2555	1	A	34,965	230	137	-	-
		B	36,476	201	171	-	-
		C	36,211	248	178	-	-
		D	37,289	313	284	-	-
		รวมร้อยละ	144,941 (100)	992 (0.68)	770 (0.50)	-	-
2555-2556	2	A	57,200	175	123	99	124
		B	57,225	180	147	113	123
		C	60,735	167	156	146	210
		D	62,627	203	223	216	190
		รวมร้อยละ	237,787 (100)	725 (0.30)	649 (0.43)	574 (0.24)	647 (0.27)
2556	3	A	33,644	166	107	-	-
		B	30,471	129	140	-	-
		C	37,006	227	178	-	-
		D	35,350	234	218	-	-
		รวมร้อยละ	136,471 (100)	756 (0.48)	643 (0.47)	-	
เฉลี่ยต่อ 1 samset และ rotation ร้อยละ			46,605 (100)	185 (0.40)	198 (0.42)	167 (0.36)	158 (0.34)
รวมทั้งหมด ร้อยละ			2,423,473 (100)	9,595 (0.40)	10,276 (0.42)	8,680 (0.36)	8,235 (0.34)

จากตารางที่ 4.7 แสดงจำนวนคนที่ไม่สามารถเชื่อมโยงได้ด้วย caseid และมีการสร้าง caseid_new ด้วยเกณฑ์ 2 (คนเดียวกันต้องมีอายุห่างกันไม่เกิน 1 ปี) ซึ่งเมื่อพิจารณาข้อมูลตัวอย่างตั้งแต่ปี พ.ศ. 2545-2556 มีจำนวนเฉลี่ย 46,605 คน เมื่อพิจารณาจำนวนคนที่อยู่ในครัวเรือนเดียวกัน มีเพศและอายุเท่ากัน และถูกสัมภาษณ์รอบที่ 1 มีจำนวนเฉลี่ย 185 คน คิดเป็นร้อยละ 0.40 นั้นหมายความว่า การใช้ caseid_new2 ที่สร้างขึ้นตามเกณฑ์ 2 มีโอกาสเกิดข้อผิดพลาดประมาณร้อยละ 0.4

บทที่ 5

บทสรุป

ในการจัดการข้อมูลตัวอย่างซ้ำรายบุคคลจากข้อมูลการสำรวจภาวะการทำงานของประชากร ตั้งแต่ปี พ.ศ. 2545-2556 นักวิจัยได้ใช้เกณฑ์การเชื่อมโยงข้อมูลตัวอย่างซ้ำรายบุคคลจาก 2 กรณี คือ 1) กรณีที่ข้อมูลตัวอย่างซ้ำรายบุคคลสามารถใช้ caseid เดิมได้ และ 2) กรณีที่ข้อมูลตัวอย่างซ้ำรายบุคคลไม่สามารถใช้ caseid เดิมได้ ตามรูปที่ 5.1 ดังนี้



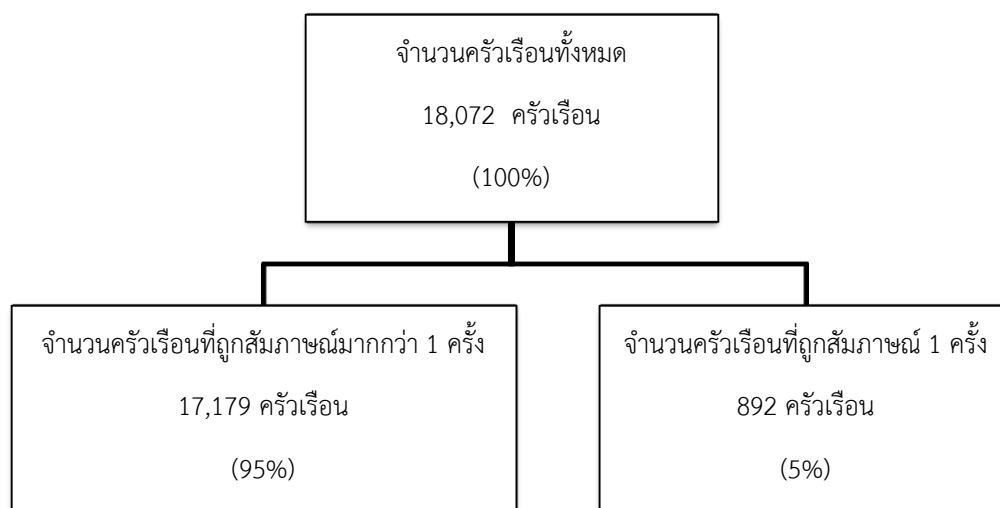
รูปที่ 5.1: การเชื่อมโยงข้อมูลตัวอย่างซ้ำรายบุคคล
จากข้อมูลการสำรวจภาวะการทำงานของประชากร

โดยเกณฑ์ที่นักวิจัยใช้ในการค้นหาข้อมูลตัวอย่างซ้ำรายบุคคล แบ่งออกเป็น 5 เกณฑ์ ได้แก่

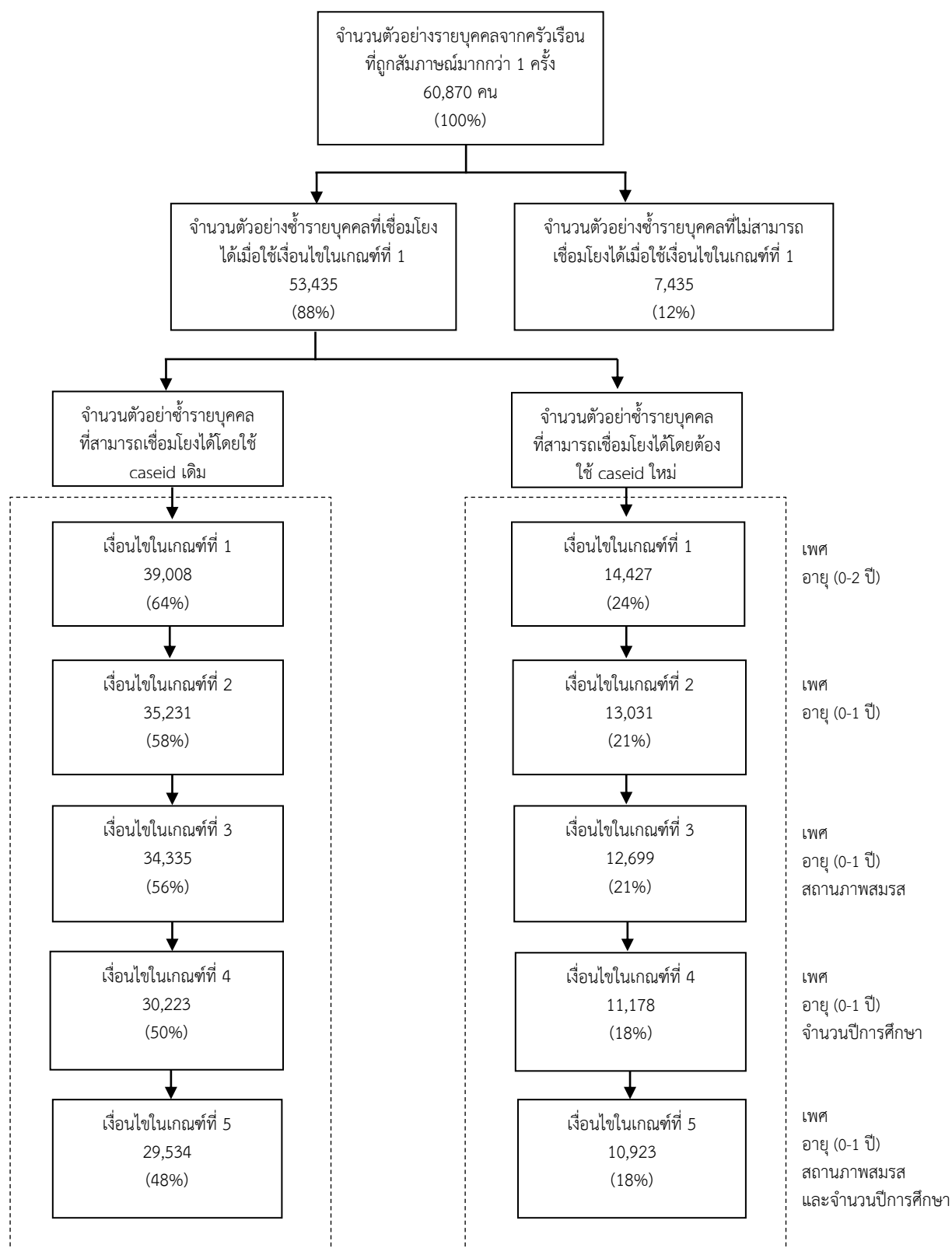
1. กลุ่มตัวอย่างซ้ำคนเดียวกันต้องมีเพศเดียวกัน และอายุไม่เกิน 2 ปี นับตั้งแต่การถูกสัมภาษณ์ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่สัมภาษณ์
2. กลุ่มตัวอย่างซ้ำคนเดียวกันต้องมีเพศเดียวกัน และอายุไม่เกิน 1 ปี นับตั้งแต่การถูกสัมภาษณ์ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่สัมภาษณ์
3. กลุ่มตัวอย่างซ้ำคนเดียวกันต้องมีเพศเดียวกัน และอายุไม่เกิน 1 ปี นับตั้งแต่การถูกสัมภาษณ์ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่สัมภาษณ์ และสถานภาพสมรสต้องเป็นไปตามเงื่อนไขของแต่ละสถานภาพ
4. กลุ่มตัวอย่างซ้ำคนเดียวกันต้องมีเพศเดียวกัน และอายุไม่เกิน 1 ปี นับตั้งแต่การถูกสัมภาษณ์ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่สัมภาษณ์ และจำนวนปีการศึกษาจะต้องมีจำนวนเท่าเดิมหรือเพิ่มขึ้นไม่เกิน 1 ปีการศึกษานับตั้งแต่การถูกสัมภาษณ์ครั้งแรกจนถึงครั้งสุดท้าย
5. กลุ่มตัวอย่างซ้ำคนเดียวกันต้องมีเพศเดียวกัน และอายุไม่เกิน 1 ปี นับตั้งแต่การถูกสัมภาษณ์ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่สัมภาษณ์ และสถานภาพสมรสเป็นไปตามเงื่อนไข และจำนวนปีการศึกษาจะต้องมีจำนวนเท่าเดิมหรือเพิ่มขึ้นไม่เกิน 1 ปีการศึกษานับตั้งแต่การถูกสัมภาษณ์ครั้งแรกจนถึงครั้งสุดท้าย

จากเกณฑ์ข้างต้น สามารถสรุปจำนวนข้อมูลกลุ่มตัวอย่างซ้ำรายบุคคลเฉลี่ยต่อ 1 rotation plan (เช่น samset A rotation 1) ได้ตามรูปที่ 5.2 และ 5.3 ซึ่งจะเห็นได้ว่า มีจำนวนครัวเรือนทั้งหมดโดยเฉลี่ย 18,072 ครัวเรือน แบ่งเป็นจำนวนครัวเรือนที่ถูกสัมภาษณ์เพียง 1 ครั้ง มีเฉลี่ย 892 ครัวเรือน คิดเป็นร้อยละ 5 ของครัวเรือนทั้งหมด และจำนวนครัวเรือนที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง มีเฉลี่ย 17,179 ครัวเรือน คิดเป็นร้อยละ 95 ของครัวเรือนทั้งหมด โดยที่ครัวเรือนที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง มีจำนวนตัวอย่างรายบุคคลทั้งหมดโดยเฉลี่ย 60,870 คน ซึ่งเมื่อใช้เกณฑ์การค้นหาข้อมูลตัวอย่างซ้ำรายบุคคลที่อยู่ในครัวเรือนที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง พบว่า หากใช้เกณฑ์ที่ 1 ซึ่งมีความรัดกุมน้อยที่สุด จะมีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ย 53,435 คน คิดเป็นร้อยละ 88 ของจำนวนคนทั้งหมด ในขณะที่หากใช้เกณฑ์ที่ 2 จะมีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ย 48,262 คน คิดเป็นร้อยละ 79 ของจำนวนคนทั้งหมด และหากใช้เกณฑ์ที่ 3 จะมีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ย 47,034 คน คิดเป็นร้อยละ 77 ของจำนวนคนทั้งหมด และใช้เกณฑ์ที่ 4 จะมีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ย 41,401 คน คิดเป็นร้อยละ 68 ของจำนวนคนทั้งหมด และเมื่อใช้เกณฑ์ที่ 5 ซึ่งรัดกุมที่สุดจะมีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ย 40,457 คน คิดเป็นร้อยละ 66 ของจำนวนคนทั้งหมด

นอกจากนี้ นักวิจัยได้ตรวจสอบโอกาสความผิดพลาดจากการเชื่อมโยงข้อมูลในกรณีที่มีสมาชิกในครัวเรือนเดียวกัน มีเพศเดียวกัน มีอายุเท่ากัน และอยู่ในรอบเดียวกันที่ถูกสัมภาษณ์ (เช่น ครัวเรือนที่มีสมาชิกเป็นฝาแฝด) ทำให้การเชื่อมโยงสมาชิกในรอบถัดไปเกิดการสลับที่ได้ โดยมีความผิดพลาดเฉลี่ยร้อยละ 0.4 ของจำนวนตัวอย่างซ้ำที่ไม่สามารถเชื่อมโยงได้โดยใช้ caseid ทั้งนี้ขึ้นอยู่กับผู้ใช้ข้อมูลที่ต้องตัดสินใจว่าจะเลือกใช้เกณฑ์ใด



รูปที่ 5.2: จำนวนข้อมูลตัวอย่างซ้ำรายครัวเรือนเฉลี่ยต่อ 1 rotation plan



รูปที่ 5.3: จำนวนข้อมูลตัวอย่างรายบุคคลเฉลี่ยต่อ 1 rotation plan

ข้อเสนอแนะ

ข้อมูลตัวอย่างข้าราชการบุคคลจากข้อมูลการสำรวจภาวะการทำงานของประชากร จะช่วยให้นักวิจัยสามารถศึกษาพฤติกรรมต่าง ๆ ด้านแรงงานที่ต้องการควบคุมลักษณะเฉพาะบุคคลได้แม่นยำขึ้น ยกตัวอย่าง เช่น งานวิจัยที่เกี่ยวข้องกับการเปลี่ยนงานของแรงงาน พลวัตของการจ้างงาน หรือผลกระทบของการเปลี่ยนแปลงระบบการกำหนดค่าจ้างขั้นต่ำ เช่นเดียวกับตัวอย่างงานวิจัยในต่างประเทศ เช่น Lemieux (2006), Autor (2011), Acemoglu and Autor (2011) และ Marquis et al. (2014)

อย่างไรก็ตาม ข้อมูลชุดนี้ก็มีข้อจำกัดที่นักวิจัยจะต้องระมัดระวังทุกครั้งที่น่าไปใช้ ประเด็นแรกคือ ตัวอย่างใด ๆ จะมีข้อมูลซ้ำได้มากที่สุดเพียง 4 ช่วงใน 2 ปี (เช่น ในปี 2545 ไตรมาสที่ 1 และ 2 และในปี 2546 ไตรมาสที่ 1 และ 2 เป็นต้น) ทำให้แบบจำลองที่สามารถประยุกต์ใช้กับข้อมูลชุดนี้จะต้องออกแบบมาสำหรับข้อมูลตัวอย่างซ้ำที่ทับซ้อนบางส่วนเป็นช่วง ๆ และอีกประเด็นหนึ่งคือ การสำรวจภาวะการทำงานของประชากรในส่วนนี้จะไม่ได้ติดตามตัวอย่างที่ย้ายที่อยู่ ทำให้ไม่สามารถใช้ข้อมูลนี้ศึกษาประเด็นที่เกี่ยวข้องกับการย้ายถิ่นของแรงงานได้

นอกจากนี้ เพื่อส่งเสริมให้มีการนำข้อมูลตัวอย่างข้าราชการบุคคลไปใช้เพื่อการวิจัย โครงการฯ หวังว่าสำนักงานสถิติแห่งชาติจะสามารถช่วยอำนวยความสะดวกให้กับนักวิจัยได้ด้วยการยินยอมให้โครงการฯ เผยแพร่ข้อมูลที่สร้างขึ้น เพราะเป็นเพียงส่วนหนึ่งไม่ใช่ทั้งหมดของข้อมูลการสำรวจภาวะการทำงานของประชากร (LFS) หรือน้อยกว่า ควรจะต้องยินยอมให้ข้อมูลที่จำเป็นต่อการสร้าง ID รายบุคคล (ข้อมูลรหัสภาค, รหัสจังหวัด, รหัสพื้นที่, psumnumber, samset, rotation, รหัสครัวเรือน และรหัสสมาชิกครัวเรือน) แก่นักวิจัยที่ต้องการใช้ประโยชน์จะได้นำข้อมูลมาประมวลผลโดยใช้ชุดคำสั่งจากโครงการฯ เพื่อนำไปศึกษาในประเด็นต่าง ๆ ต่อไป

เอกสารอ้างอิง

- Acemoglu, D. and Autor, D. (2011) “Chapter 12 - Skills, Tasks and Technologies: Implications for Employment and Earnings” *Handbook of Labor Economics*, 4: 1043 – 1171.
- Autor, D. (2011) “The Polarization of Job Opportunities in the U.S. Labor Market” *Community Investments*, 23(2): 11-16.
- Fairlie, R., 2014, Recent Trends in Ethnic and Racial Business Ownership. University of California.
- Hoynes, H., Miller, D., L., Schaller. 2012. Who Suffers During Recessions?. *Journal of Economic Perspectives*. Vol. 26, No.3, p.27-48.
- Lemieux, T. (2006) “Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill?” *American Economic Review*, 96(3): 461-498.
- Madrian, B., C., Lefgren, L., J. 2000. An Approach to Longitudinally matching Current Population Survey (CPS) respondents. *Journal of Economic and Social Measurement*. pp. 31-62.
- Marquis M. H., Trehan B., and Tantivong W. (2014) “The wage premium puzzle and the quality of human capital” *International Review of Economics & Finance*, 33: 100-110.
- Mincer, Jacob. "Schooling." *Experience and Earnings*, New York (1974).
- Neumark, D., Wascher, W. 1992. Employment Effects of Minimum and Subminimum Wages: Panel Data on State Minimum Wage Laws. *Industrial and Labor Relations Review*. Vol. 46, No.1(October 1992)
- Warunsiri, S., Mcnown, R. 2010. The Return to Education in Thailand: A Pseudo-Panel Approach. *World Development*. Vol. 38, No. 11, pp. 1616-1625

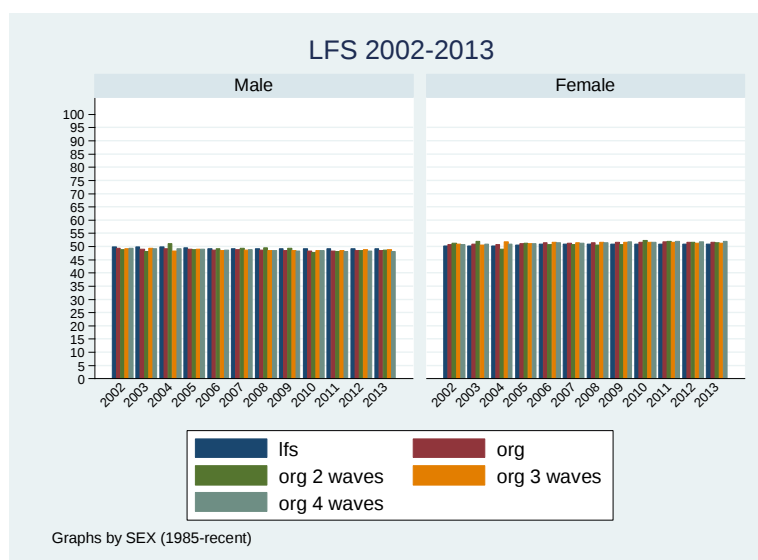
ภาคผนวก ก.

สถิติเบื้องต้นของข้อมูลรายบุคคลแบบตัวอย่างซ้ำจากข้อมูลภาวะการทำงานของประชากร

ค่าสถิติเบื้องต้นของข้อมูลรายบุคคลแบบตัวอย่างซ้ำ (ORG) จากข้อมูลภาวะการทำงานของประชากร (LFS)

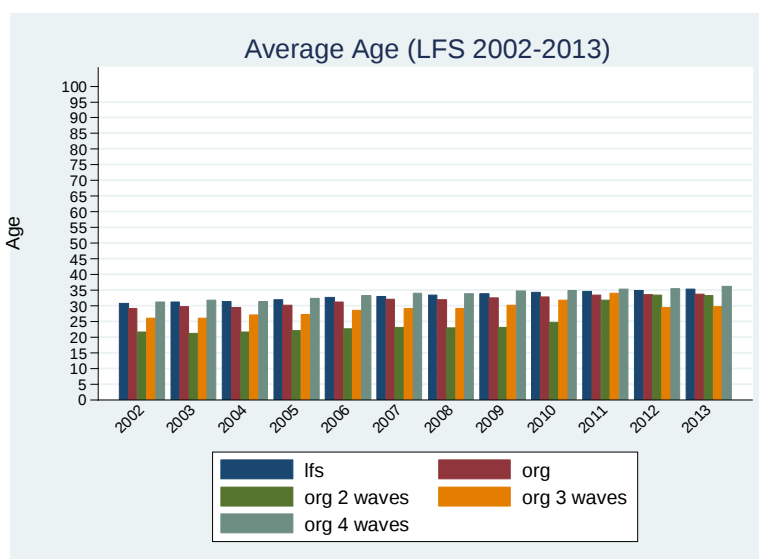
ในบทนี้เป็นการนำเสนอค่าสถิติเบื้องต้นจากข้อมูลการสำรวจภาวะการทำงานของประชากร (LFS) และข้อมูลรายบุคคลแบบตัวอย่างซ้ำจากข้อมูลการสำรวจภาวะการทำงานของประชากร (ORG) ซึ่งจากค่าสถิติต่าง ๆ ที่นำเสนอในส่วนนี้ พบว่า ข้อมูล ORG มีแนวโน้มใกล้เคียงกับข้อมูล LFS แสดงให้เห็นว่าข้อมูล ORG มีขนาดมากพอและยังคงเป็นตัวแทนที่ดีที่สามารถนำมาวิเคราะห์ในประเด็นต่าง ๆ ได้เช่นเดียวกับข้อมูล LFS

เมื่อพิจารณาร้อยละของจำนวนกลุ่มตัวอย่างที่แบ่งตามเพศ พบว่า ทั้งข้อมูล LFS และ ORG มีร้อยละของเพศหญิงสูงกว่าเพศชาย โดยเฉลี่ยคิดเป็นเพศชายร้อยละ 49 และเพศหญิงร้อยละ 51 นั่นคือ สัดส่วนของเพศชายและเพศหญิงจากข้อมูล ORG มีค่าใกล้เคียงกับข้อมูล LFS ตามรูปที่ 1



รูปที่ 1: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่แบ่งตามเพศ

เมื่อพิจารณาอายุเฉลี่ยของข้อมูล LFS, ข้อมูลตัวอย่างซ้ำ ORG ที่กลุ่มตัวอย่างซ้ำถูกสัมภาษณ์มากกว่า 1 ครั้งขึ้นไป และข้อมูลกลุ่มตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์ 2 ครั้ง, 3 ครั้ง และ 4 ครั้ง ตามลำดับจากรูปที่ 2 พบว่า ข้อมูลกลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 2 ครั้งมีอายุเฉลี่ยต่ำกว่ากลุ่มตัวอย่างจากข้อมูลชุดอื่น ๆ ขณะที่ข้อมูลกลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 4 ครั้งมีอายุเฉลี่ยสูงที่สุด



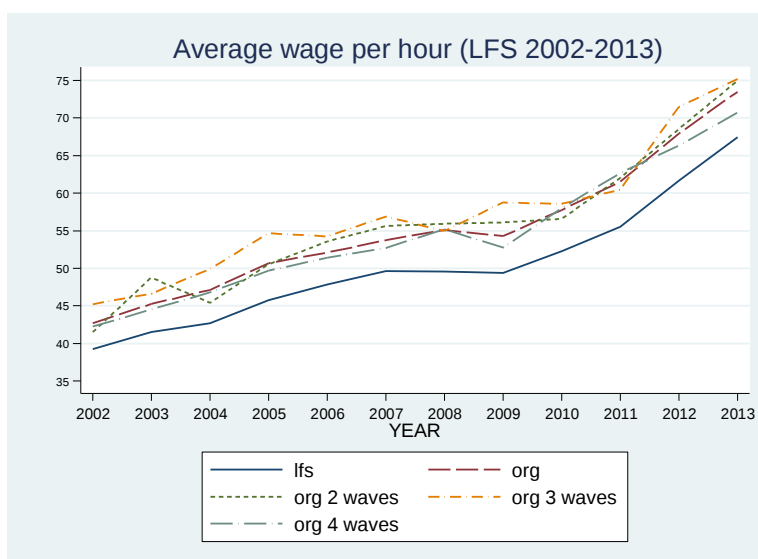
รูปที่ 2: อายุเฉลี่ยของกลุ่มตัวอย่างจากข้อมูล LFS และ ORG

ตารางที่ 1: ค่าสถิติพื้นฐานของตัวแปรอายุจากข้อมูล LFS และ ORG

<i>variable</i>	<i>observation</i>	<i>mean</i>	<i>standard deviation</i>	<i>min (age average)</i>	<i>max (age average)</i>
age_lfs	12	33.05	1.5367	30.7314	35.2537
age_org	12	31.59	1.6428	29.1525	33.6164
age_org_2w	12	25.07	4.7488	21.0908	33.3738
age_org_3w	12	28.98	2.3082	26.0713	33.9422
age_org_4w	12	33.67	1.6994	31.2168	36.1779

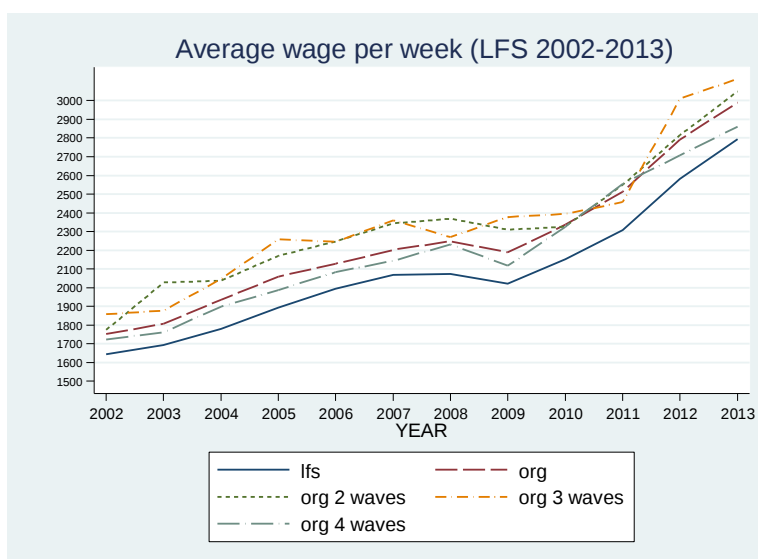
เมื่อทำการเปรียบเทียบอายุเฉลี่ยของข้อมูล ORG แต่ละชุดกับข้อมูล LFS ซึ่งมีผลลัพธ์ตามตารางที่ 1 พบว่า อายุเฉลี่ยของกลุ่มตัวอย่างในแต่ละชุดข้อมูลมีค่าแตกต่างกัน โดยมีอายุเฉลี่ยที่เรียงจากน้อยไปหามาก ดังนี้ อายุเฉลี่ยของข้อมูล ORG 2 waves (ประมาณ 25 ปี), ข้อมูล ORG 3 waves (ประมาณ 30 ปี), ข้อมูล LFS (ประมาณ 33 ปี), ข้อมูล ORG (ประมาณ 32 ปี), และข้อมูล ORG 4 waves (ประมาณ 34 ปี) ตามลำดับ ซึ่งจะเห็นได้ว่าอายุเฉลี่ยของตัวอย่างจากข้อมูล ORG มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด

เมื่อพิจารณาค่าจ้างเฉลี่ยรายชั่วโมงของข้อมูล LFS และ ORG จากรูปที่ 3 พบว่า ข้อมูลตัวอย่างซ้ำที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง, ถูกสัมภาษณ์ 2 ครั้ง, 3 ครั้ง และ 4 ครั้ง มีแนวโน้มและทิศทางเช่นเดียวกับข้อมูลจาก LFS โดยที่ข้อมูลตัวอย่างซ้ำที่ถูกสัมภาษณ์ 4 ครั้งมีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด



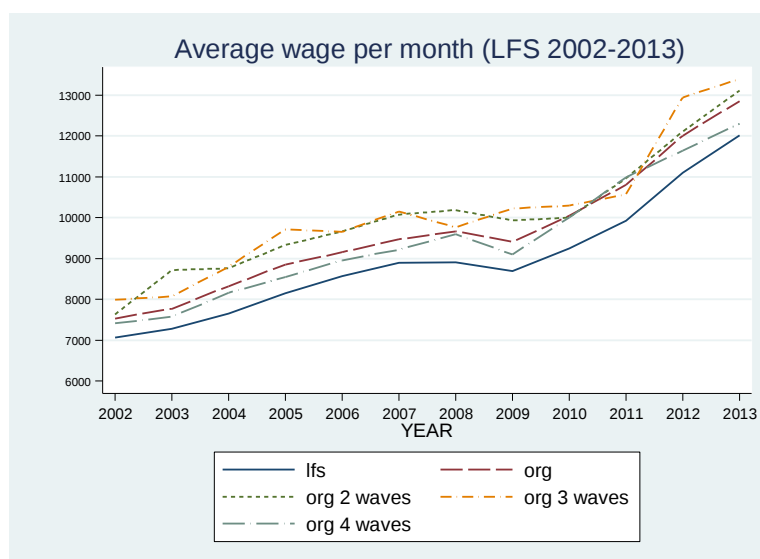
รูปที่ 3: ค่าจ้างเฉลี่ยรายชั่วโมงของข้อมูล LFS และ ORG

เมื่อพิจารณาค่าจ้างเฉลี่ยรายสัปดาห์ของข้อมูล LFS และ ORG จากรูปที่ 4 พบว่า ข้อมูลตัวอย่างซ้ำที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง, ถูกสัมภาษณ์ 2 ครั้ง, 3 ครั้ง และ 4 ครั้ง มีแนวโน้มและทิศทางเช่นเดียวกับข้อมูล LFS โดยที่ข้อมูลตัวอย่างซ้ำที่ถูกสัมภาษณ์ 4 ครั้งมีค่าใกล้เคียงกับข้อมูล LFS มากที่สุดเช่นกัน



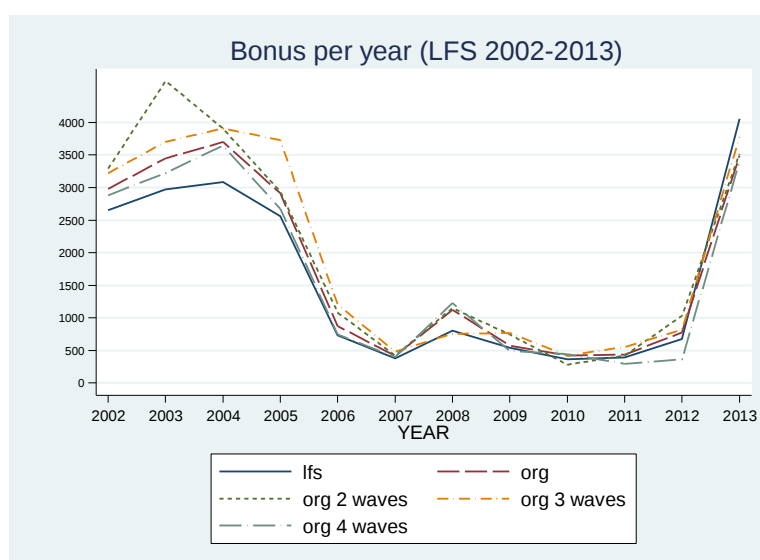
รูปที่ 4: ค่าจ้างเฉลี่ยรายสัปดาห์จากข้อมูล LFS และ ORG

เมื่อพิจารณาค่าจ้างเฉลี่ยรายเดือนของข้อมูล LFS และ ORG จากรูปที่ 5 พบว่า ข้อมูลตัวอย่างซ้ำที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง, ถูกสัมภาษณ์ 2 ครั้ง, 3 ครั้ง และ 4 ครั้ง มีแนวโน้มและทิศทางเช่นเดียวกับข้อมูล LFS โดยที่ข้อมูลตัวอย่างซ้ำที่ถูกสัมภาษณ์ 4 ครั้งยังคงมีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด



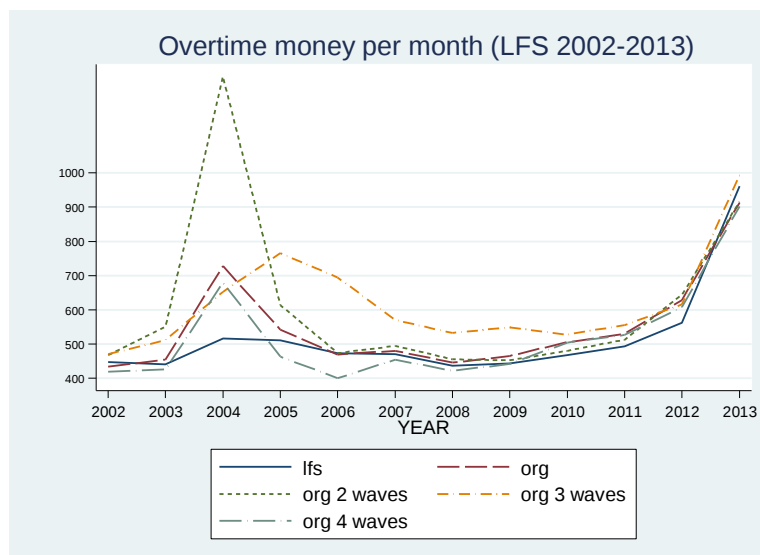
รูปที่ 5: ค่าจ้างเฉลี่ยรายเดือนจากข้อมูล LFS และ ORG

เมื่อพิจารณาค่าโบนัสเฉลี่ยรายปีของข้อมูล LFS และ ORG จากรูปที่ 6 พบว่า ข้อมูลตัวอย่างซ้ำที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง, ถูกสัมภาษณ์ 2 ครั้ง, 3 ครั้ง และ 4 ครั้ง มีแนวโน้มและทิศทางเช่นเดียวกับข้อมูลจาก LFS โดยที่ข้อมูลตัวอย่างซ้ำที่ถูกสัมภาษณ์ 4 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS ซึ่งจะเห็นได้ว่าคุณค่าโบนัสเฉลี่ยรายปีในช่วง 2006-2012 มีค่าต่ำลง และเพิ่มสูงขึ้นอีกครั้งในปี 2013



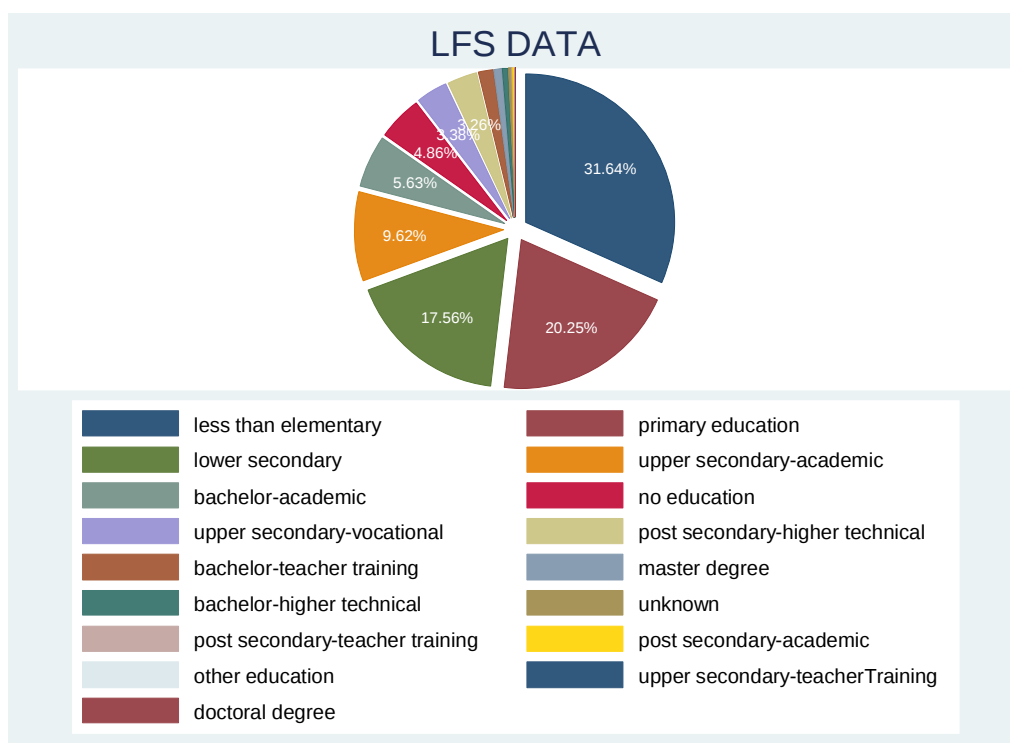
รูปที่ 6: ค่าโบนัสเฉลี่ยรายปีจากข้อมูล LFS และ ORG

เมื่อพิจารณาค่าโอทีหรือเงินล่วงเวลาเฉลี่ยรายเดือนของข้อมูล LFS และ ORG จากรูปที่ 7 พบว่า ข้อมูลตัวอย่างซ้ำที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง, ถูกสัมภาษณ์ 2 ครั้ง, 3 ครั้ง และ 4 ครั้ง มีแนวโน้มสูงกว่าข้อมูลจาก LFS โดยเฉพาะอย่างยิ่งในปี 2004 ที่ค่าโอทีและค่าล่วงเวลามีค่าเฉลี่ยสูงที่สุด



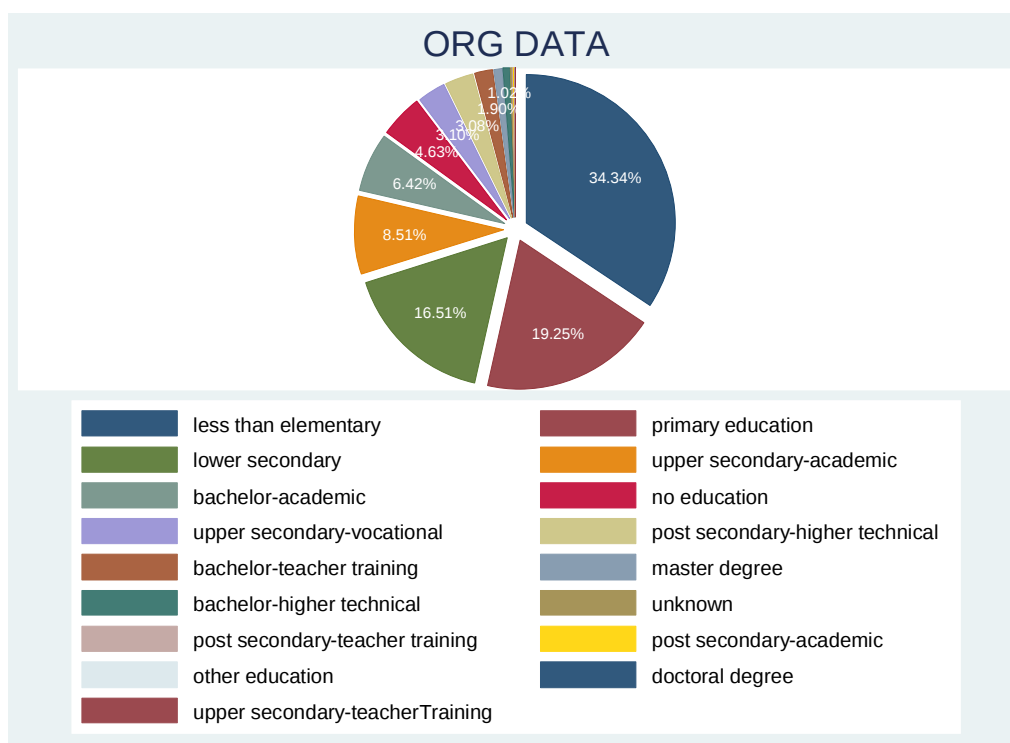
รูปที่ 7: ค่าโอทีเฉลี่ยรายเดือนจากข้อมูล LFS และ ORG

เมื่อพิจารณาร้อยละของผู้ที่จบการศึกษาในระดับต่าง ๆ จากข้อมูล LFS, ข้อมูลตัวอย่างซ้ำ ORG ที่กลุ่มตัวอย่างซ้ำถูกสัมภาษณ์มากกว่า 1 ครั้งขึ้นไป และข้อมูลกลุ่มตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์ 2 ครั้ง, 3 ครั้ง และ 4 ครั้ง ตามลำดับ ซึ่งปรากฏตามรูปที่ 8 – 12 ดังต่อไปนี้



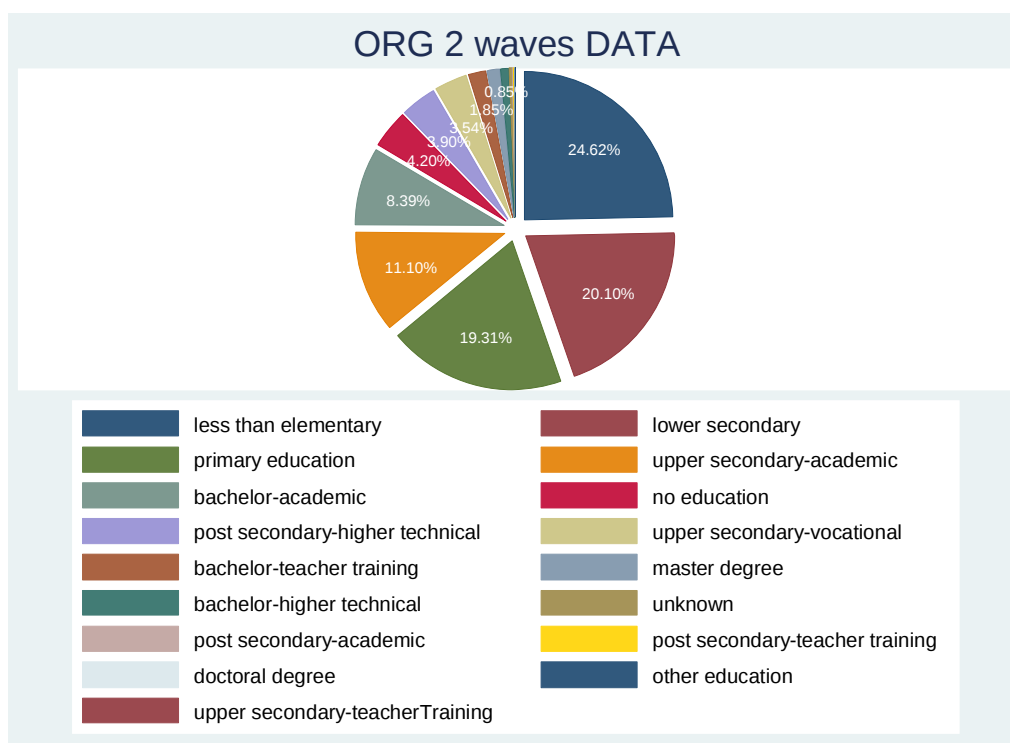
รูปที่ 8: ร้อยละโดยเฉลี่ยตั้งแต่ปี 2002-2013 ของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS
แบ่งตามระดับการศึกษา

จากรูปที่ 8 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่แบ่งตามการจบระดับการศึกษาสูงสุดจากข้อมูล LFS ตั้งแต่ปี 2002-2013 พบว่า โดยเฉลี่ยคนที่จบการศึกษาต่ำกว่าชั้นประถมศึกษา มีจำนวนสูงที่สุดคิดเป็น ร้อยละ 31.64 ตามด้วยคนที่จบการศึกษาระดับประถมศึกษา ร้อยละ 20.25, คนที่จบการศึกษาระดับมัธยมศึกษาตอนต้น ร้อยละ 17.56, คนที่จบการศึกษามัธยมศึกษาตอนปลาย ร้อยละ 9.62 และคนที่จบการศึกษาปริญญาตรี (สายวิชาการ) ร้อยละ 5.63 ตามลำดับ



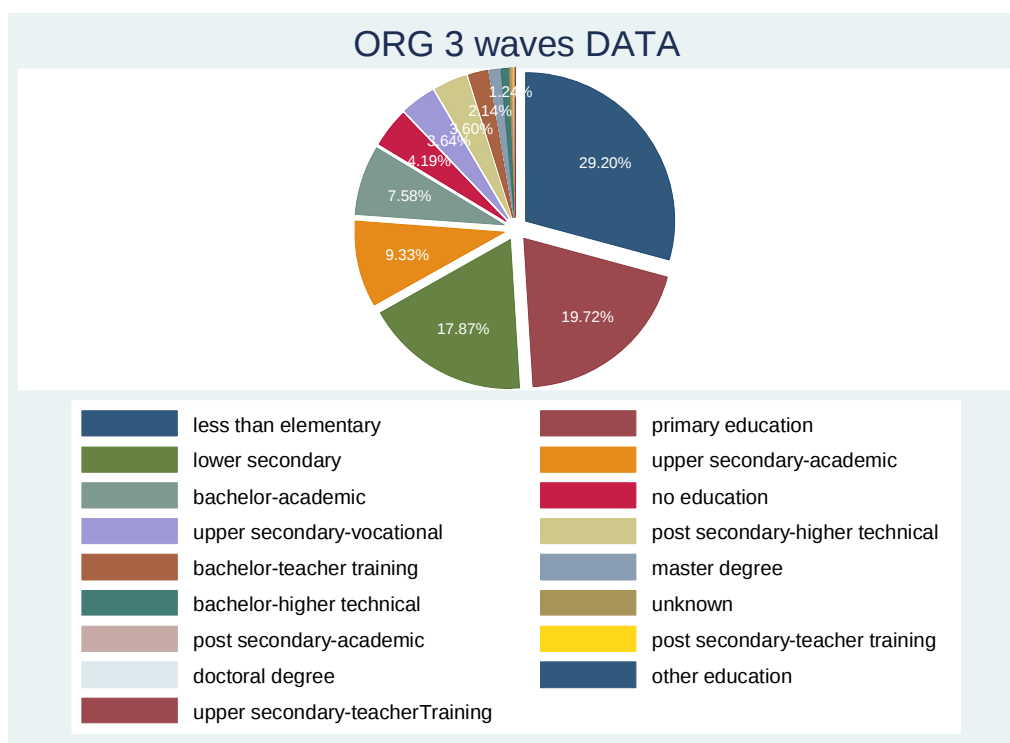
รูปที่ 9: ร้อยละโดยเฉลี่ยตั้งแต่ปี 2002-2013 ของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG
แบ่งตามระดับการศึกษา

จากรูปที่ 9 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่แบ่งตามการจบระดับการศึกษาสูงสุดจากข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์มากกว่า 1 ครั้งขึ้นไป ตั้งแต่ปี 2002-2013 พบว่า โดยเฉลี่ยคนที่จบการศึกษาต่ำกว่าชั้นประถมศึกษา มีจำนวนสูงที่สุดคิดเป็นร้อยละ 34.34 ตามด้วยคนที่จบการศึกษาระดับประถมศึกษา ร้อยละ 19.25, คนที่จบการศึกษาระดับมัธยมศึกษาตอนต้น ร้อยละ 16.51, คนที่จบการศึกษามัธยมศึกษาตอนปลาย ร้อยละ 8.51 และคนที่จบการศึกษาปริญญาตรี (สายวิชาการ) ร้อยละ 6.42 ตามลำดับ ซึ่งมีลำดับร้อยละของคนจบการศึกษาในระดับต่าง ๆ เหมือนกันกับข้อมูล LFS แต่จะเห็นว่า ร้อยละของคนจบการศึกษาปริญญาตรี (สายวิชาการ) จากข้อมูล ORG มีค่าสูงกว่าร้อยละของคนจบการศึกษาปริญญาตรี (สายวิชาการ) จากข้อมูล LFS



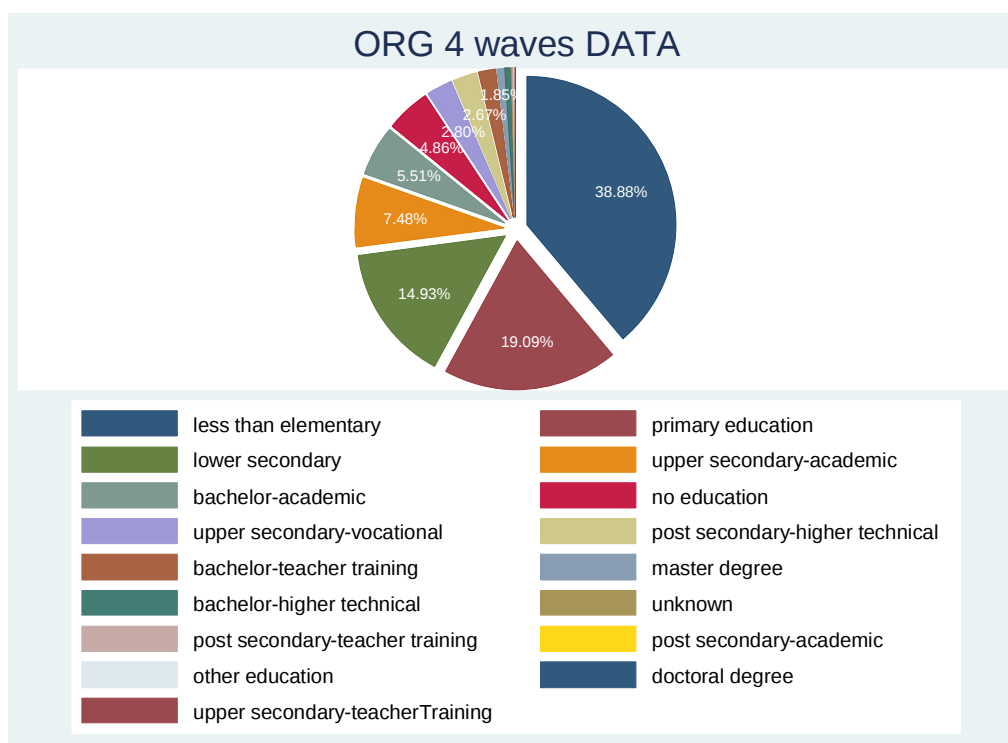
รูปที่ 10: ร้อยละโดยเฉลี่ยตั้งแต่ปี 2002-2013 ของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 2 waves แบ่งตามระดับการศึกษา

จากรูปที่ 10 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่แบ่งตามการจบระดับการศึกษาสูงสุดจากข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์ 2 ครั้ง ตั้งแต่ปี 2002-2013 พบว่า โดยเฉลี่ยคนที่จบการศึกษาต่ำกว่าชั้นประถมศึกษา มีจำนวนสูงที่สุดคิดเป็นร้อยละ 24.62 ตามด้วยคนที่จบการศึกษาระดับประถมศึกษา ร้อยละ 20.10, คนที่จบการศึกษาระดับมัธยมศึกษาตอนต้น ร้อยละ 18.31, คนที่จบการศึกษามัธยมศึกษาตอนปลาย ร้อยละ 11.10 และคนที่จบการศึกษาปริญญาตรี (สายวิชาการ) ร้อยละ 8.39 ตามลำดับ ซึ่งมีลำดับร้อยละของคนที่จบการศึกษาในระดับต่าง ๆ เหมือนกันกับข้อมูล LFS แต่จะเห็นว่า ร้อยละของคนที่จบการศึกษาต่ำกว่าระดับประถมศึกษาจากข้อมูล LFS มีค่าสูงกว่าร้อยละของคนที่จบการศึกษาต่ำกว่าระดับประถมศึกษาจากข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์ 2 ครั้ง



รูปที่ 11: ร้อยละโดยเฉลี่ยตั้งแต่ปี 2002-2013 ของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 3 waves แบ่งตามระดับการศึกษา

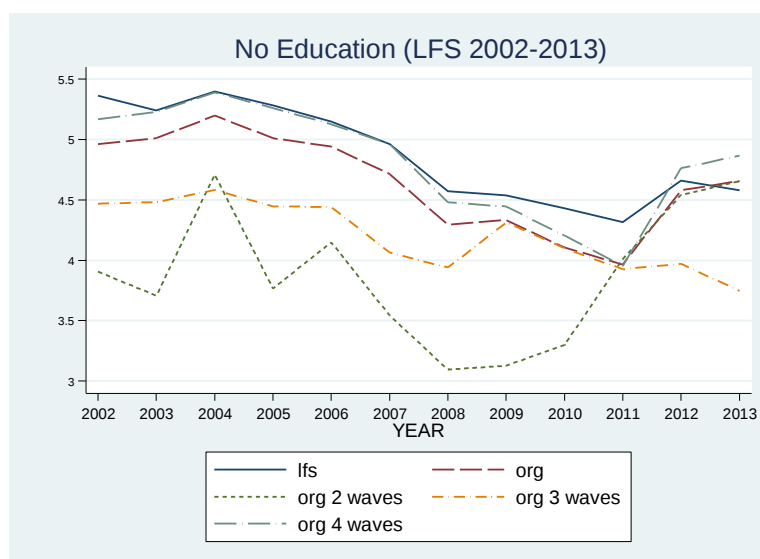
จากรูปที่ 11 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่แบ่งตามการจบระดับการศึกษาสูงสุดจากข้อมูลตัวอย่างเข้า ORG ที่ถูกสัมภาษณ์ 3 ครั้ง ตั้งแต่ปี 2002-2013 พบว่า โดยเฉลี่ยคนที่จบการศึกษาต่ำกว่าชั้นประถมศึกษา มีจำนวนสูงที่สุดคิดเป็นร้อยละ 29.20 ตามด้วยคนที่จบการศึกษาระดับประถมศึกษา ร้อยละ 19.72, คนที่จบการศึกษาระดับมัธยมศึกษาตอนต้น ร้อยละ 17.87, คนที่จบการศึกษามัธยมศึกษาตอนปลาย ร้อยละ 9.33 และคนที่จบการศึกษาปริญญาตรี (สายวิชาการ) ร้อยละ 7.58 ตามลำดับ ซึ่งมีลำดับร้อยละของคนที่จบการศึกษาในระดับต่าง ๆ เหมือนกับกับข้อมูล LFS



รูปที่ 12: ร้อยละโดยเฉลี่ยตั้งแต่ปี 2002-2013 ของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 4 waves แบ่งตามระดับการศึกษา

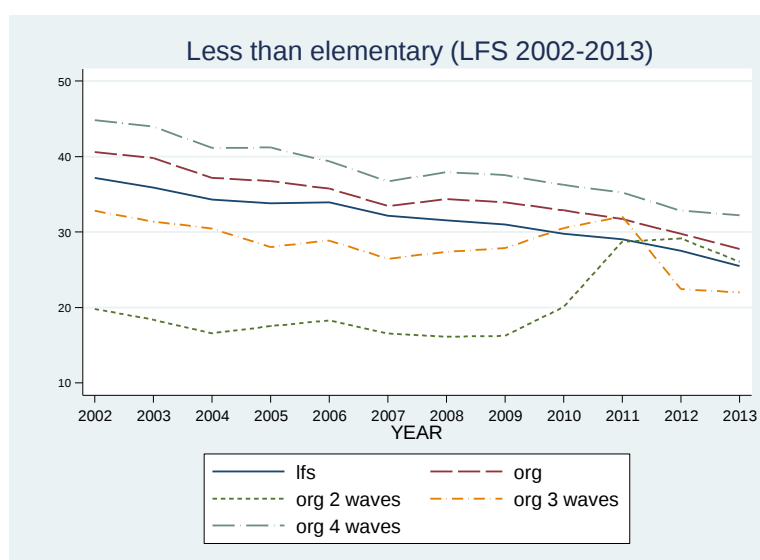
จากรูปที่ 12 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่แบ่งตามการจบระดับการศึกษาสูงสุดจากข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์ 4 ครั้ง ตั้งแต่ปี 2002-2013 พบว่า โดยเฉลี่ยคนที่จบการศึกษาต่ำกว่าชั้นประถมศึกษา มีจำนวนสูงที่สุดคิดเป็นร้อยละ 38.88 ตามด้วยคนที่จบการศึกษาระดับประถมศึกษา ร้อยละ 19.09 คนที่จบการศึกษาระดับมัธยมศึกษาตอนต้น ร้อยละ 14.93, คนที่จบการศึกษามัธยมศึกษาตอนปลาย ร้อยละ 7.48 และคนที่จบการศึกษาปริญญาตรี (สายวิชาการ) ร้อยละ 5.51 ตามลำดับ ซึ่งมีลำดับร้อยละของคนที่จบการศึกษาในระดับต่าง ๆ เหมือนกันกับข้อมูล LFS และจะเห็นได้ว่าร้อยละของคนที่จบการศึกษาตั้งแต่ระดับประถมศึกษาจนถึงระดับมัธยมศึกษาตอนปลายจากข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์ 4 ครั้งมีค่าสูงกว่าร้อยละของคนที่จบการศึกษาตั้งแต่ระดับประถมศึกษาจนถึงระดับมัธยมศึกษาตอนปลายจากข้อมูล LFS

ในส่วนถัดไปจะเป็นการเปรียบเทียบร้อยละโดยเฉลี่ยต่อปีของแต่ละชุดข้อมูลที่แบ่งตามระดับการศึกษา ซึ่งสามารถแสดงได้ดังต่อไปนี้



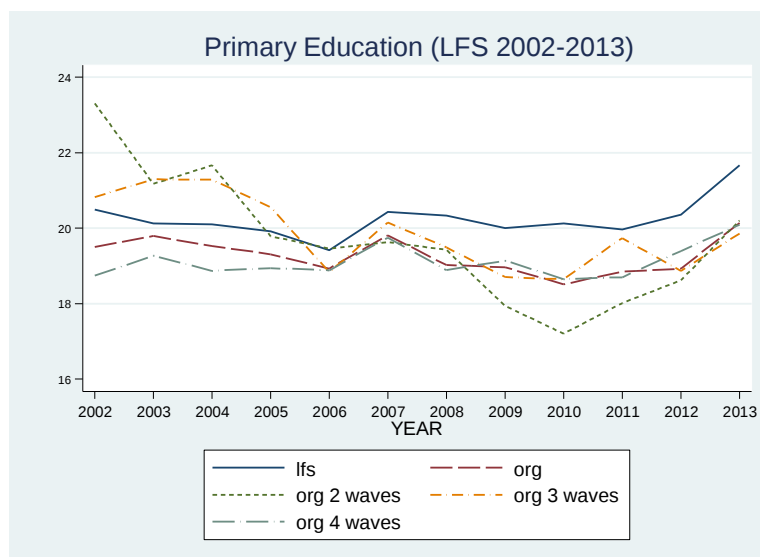
รูปที่ 13: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่ไม่มีการศึกษา

จากรูปที่ 13 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่ไม่มีการศึกษา พบว่า ทั้งข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มลดลง ยกเว้นข้อมูล ORG ตัวอย่างซ้ำที่ถูกสัมภาษณ์ 2 ครั้ง ที่มีแนวโน้มสูงขึ้น ส่วนแนวโน้มข้อมูลตัวอย่างซ้ำที่ลดลงแสดงว่า จำนวนกลุ่มตัวอย่างได้รับการศึกษาในระดับที่สูงขึ้น โดยที่ข้อมูล ORG ซึ่งมีตัวอย่างซ้ำถูกสัมภาษณ์ 4 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS มากกว่าข้อมูล ORG ชุดอื่น ๆ



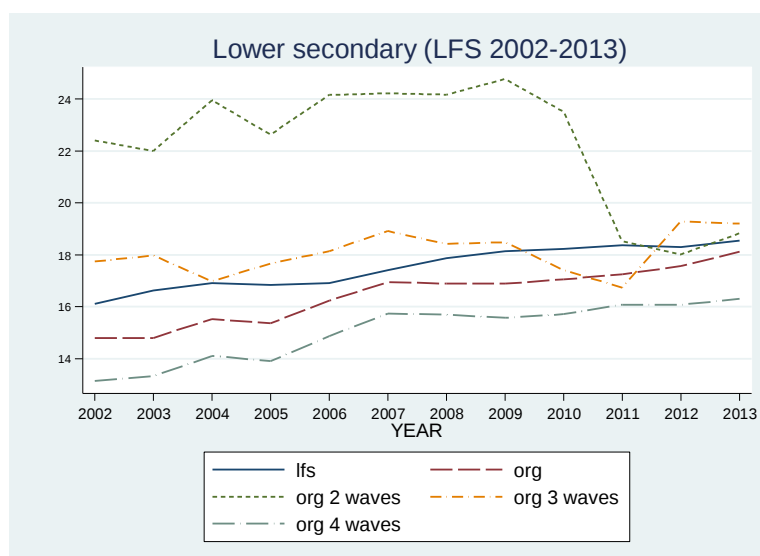
รูปที่ 14: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาต่ำกว่าชั้นประถมศึกษา

รูปที่ 14 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่จบการศึกษาต่ำกว่าชั้นประถมศึกษา พบว่า ทั้งข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มลดลงเช่นกัน แสดงให้เห็นว่าเมื่อเวลาผ่านไป จำนวนกลุ่มตัวอย่างได้รับการศึกษาในระดับที่สูงกว่าระดับประถมศึกษาเพิ่มขึ้น โดยที่ข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 3 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด



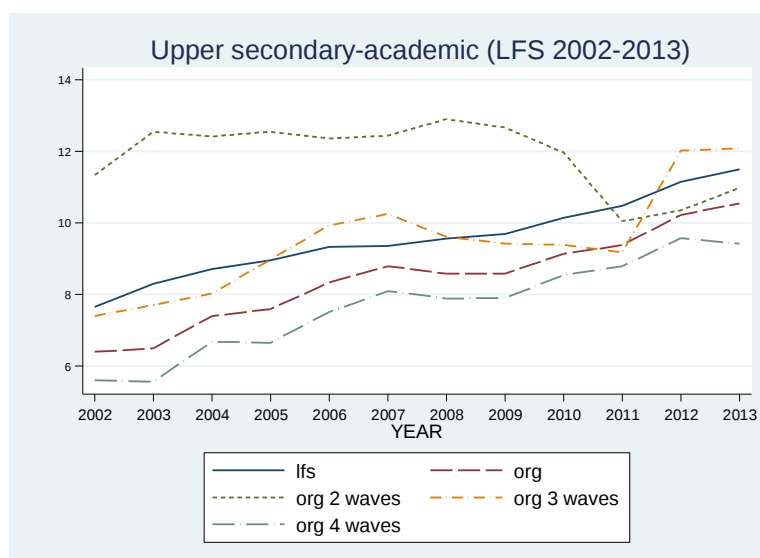
รูปที่ 15: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาชั้นประถมศึกษา

รูปที่ 15 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่จบการศึกษาชั้นประถมศึกษา พบว่า ทั้งข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มเพิ่มขึ้น แสดงว่าเมื่อเวลาผ่านไป จำนวนกลุ่มตัวอย่างที่จบการศึกษาในระดับประถมศึกษาจะเพิ่มขึ้น โดยที่ข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 3 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด



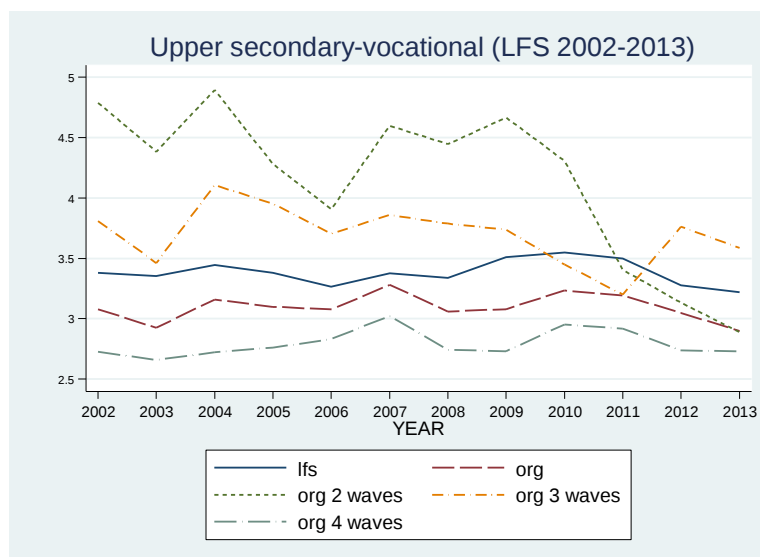
รูปที่ 16: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาชั้นมัธยมศึกษาตอนต้น

จากรูปที่ 16 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่จบการศึกษาชั้นมัธยมศึกษาตอนต้น พบว่า ทั้งข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มเพิ่มขึ้น นั่นคือ เมื่อเวลาผ่านไป จำนวนกลุ่มตัวอย่างที่จบการศึกษาในระดับมัธยมศึกษาตอนต้นเพิ่มขึ้น โดยที่ข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 3 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด



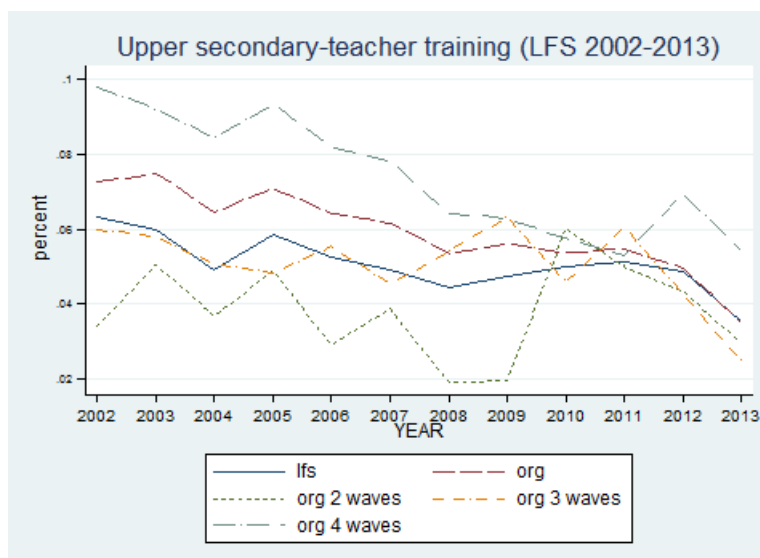
รูปที่ 17: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาชั้นมัธยมศึกษาตอนปลาย

จากรูปที่ 17 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่จบการศึกษาชั้นมัธยมศึกษาตอนปลาย พบว่า ทั้งข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มเพิ่มขึ้น นั่นคือ เมื่อเวลาผ่านไป จำนวนกลุ่มตัวอย่างที่จบการศึกษาในระดับมัธยมศึกษาตอนปลายเพิ่มขึ้น โดยที่ข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 3 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด



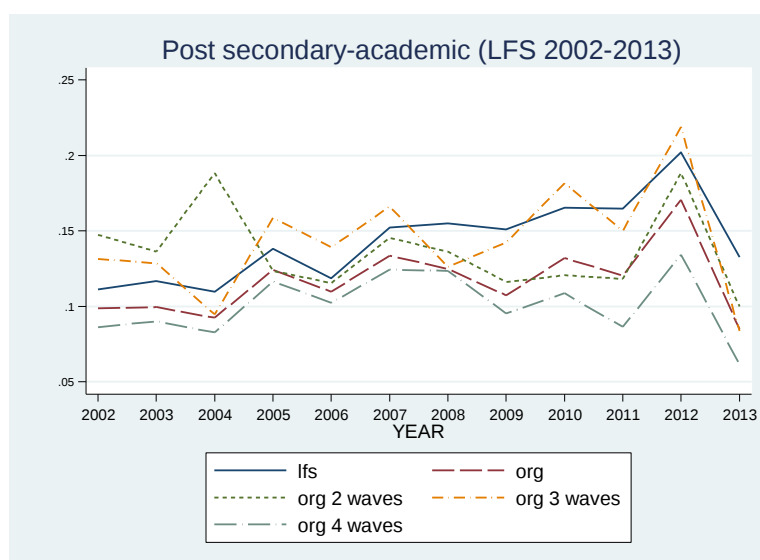
รูปที่ 18: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาชั้นอาชีวศึกษาระดับ ปวช. (เทียบกับมัธยมศึกษาตอนปลายสายสามัญ)

จากรูปที่ 18 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่จบการศึกษาชั้นอาชีวศึกษาระดับ ปวช. พบว่า ทั้งข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มลดลง นั่นคือ จำนวนกลุ่มตัวอย่างที่จบการศึกษาในระดับอาชีวศึกษา ปวช. มีจำนวนลดลง โดยที่ข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 3 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด



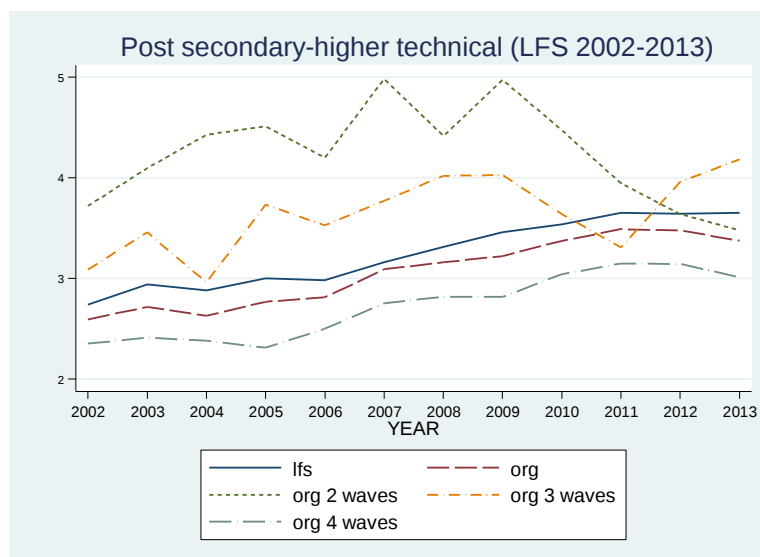
รูปที่ 19: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาสายศึกษาศาสตร์ (เทียบเท่ามัธยมศึกษาตอนปลายสายสามัญ)

จากรูปที่ 19 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่จบการศึกษาชั้นมัธยมศึกษาตอนปลายสายศึกษาศาสตร์ พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มลดลง นั่นคือ จำนวนกลุ่มตัวอย่างที่จบการศึกษาในระดับมัธยมศึกษาตอนปลายสายศึกษาศาสตร์ลดลง โดยที่ข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 3 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด



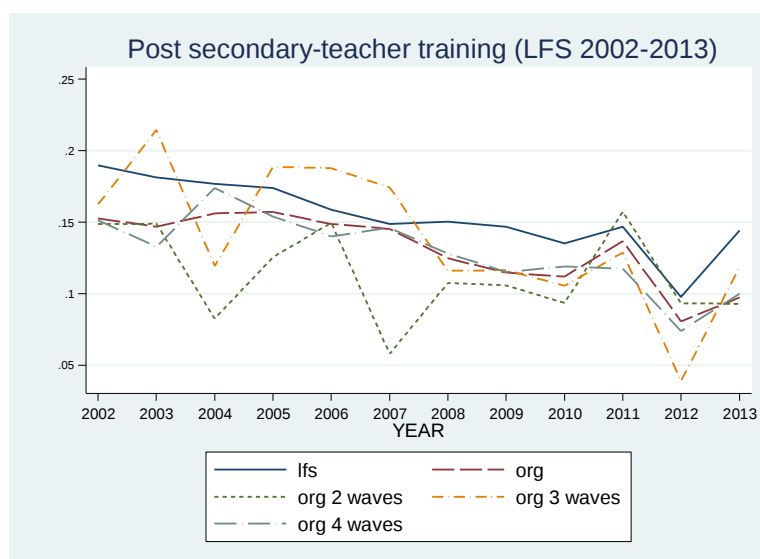
รูปที่ 20: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาชั้นอนุปริญญา (สายสามัญ)

จากรูปที่ 20 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่จบการศึกษาชั้นอนุปริญญาสายสามัญ พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มเพิ่มขึ้นตั้งแต่ปี 2002-2012 และลดลงในปี 2013 นั่นคือ จำนวนกลุ่มตัวอย่างที่จบการศึกษาในระดับอนุปริญญาสายสามัญเพิ่มขึ้น แต่กลับมีแนวโน้มลดลงในปี 2013 โดยที่ข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 3 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด



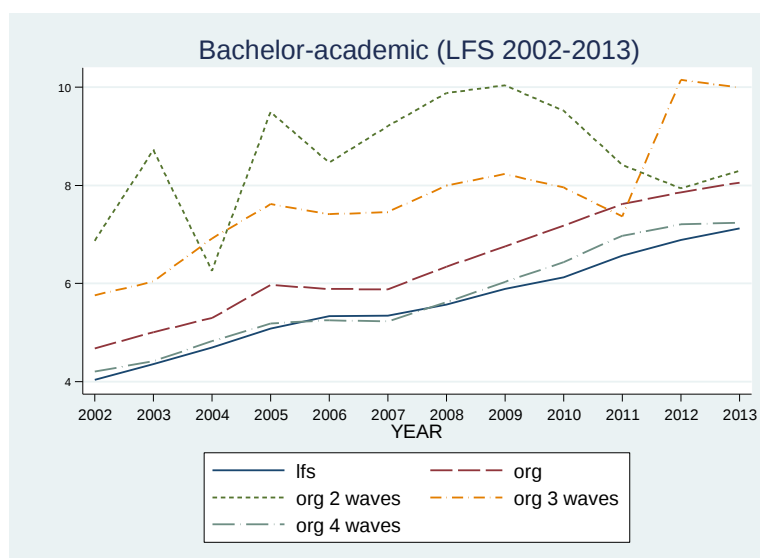
รูปที่ 21: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาชั้นอนุปริญญา (สายอาชีวศึกษา ระดับ ปวส.)

จากรูปที่ 21 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่จบการศึกษาชั้นอนุปริญญาสายอาชีวศึกษา ระดับ ปวส. พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มเพิ่มขึ้น นั่นคือ จำนวนกลุ่มตัวอย่างที่จบการศึกษาในระดับอนุปริญญาสายอาชีวศึกษา ระดับ ปวส. เพิ่มขึ้น โดยที่ข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 4 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด



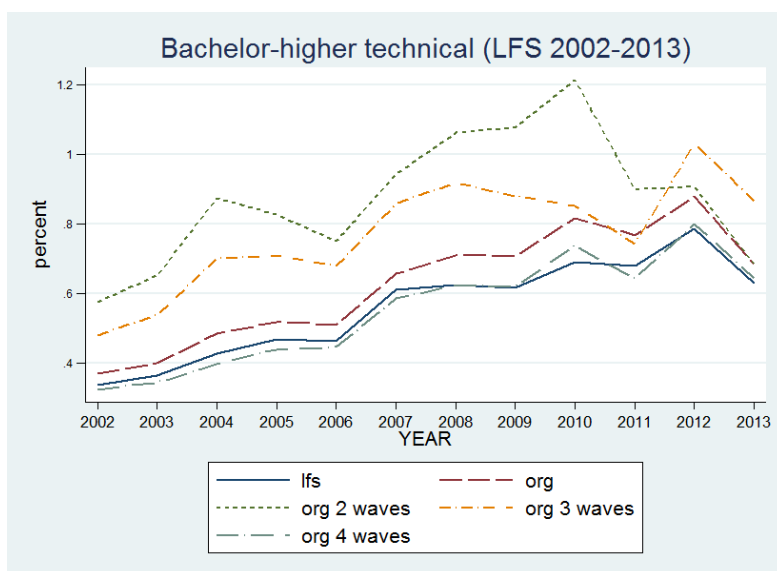
รูปที่ 22: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG
ที่จบการศึกษาชั้นอนุปริญญา (สายศึกษาศาสตร์)

จากรูปที่ 22 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่จบการศึกษาชั้นอนุปริญญาสายศึกษาศาสตร์พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มลดลง นั่นคือ จำนวนกลุ่มตัวอย่างที่จบการศึกษาในระดับอนุปริญญาสายศึกษาศาสตร์ลดลง โดยที่ข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 4 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด



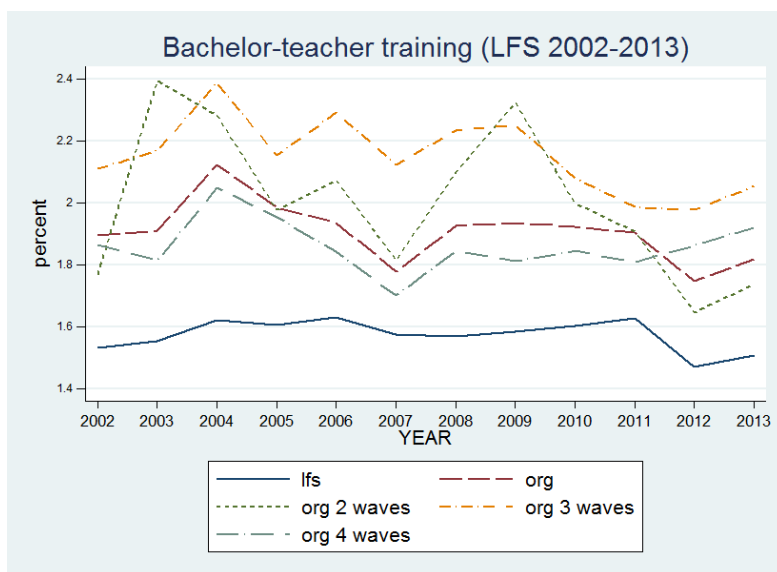
รูปที่ 23: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG
ที่จบการศึกษาระดับปริญญาตรี (สายวิชาการ)

รูปที่ 23 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่จบการศึกษาระดับปริญญาตรีสายวิชาการ พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มเพิ่มขึ้น นั่นคือ จำนวนกลุ่มตัวอย่างที่จบการศึกษาในระดับปริญญาตรีสายวิชาการเพิ่มขึ้น โดยที่ข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 4 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด



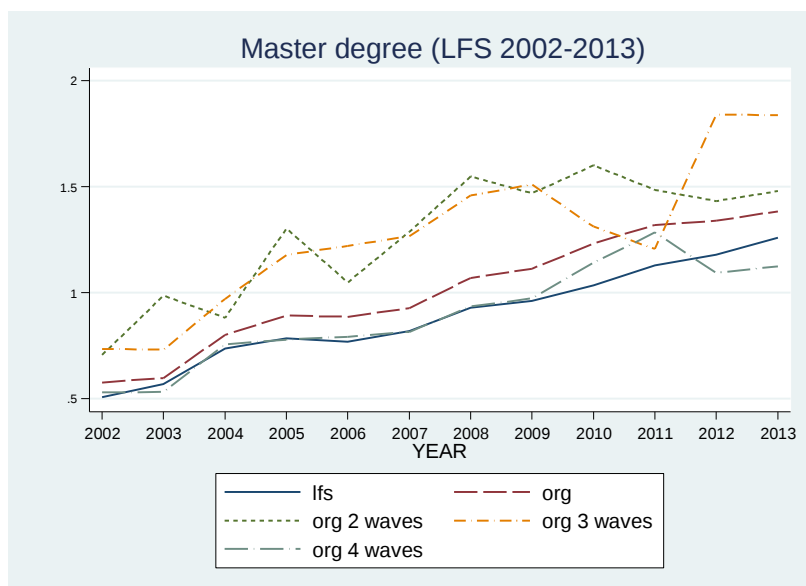
รูปที่ 24: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาระดับปริญญาตรี (สายอาชีพศึกษา)

จากรูปที่ 24 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่จบการศึกษาระดับปริญญาตรีสายอาชีพศึกษา พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มเพิ่มขึ้นสำหรับปี 2002-2012 แต่ลดลงในปี 2013 นั่นคือ จำนวนกลุ่มตัวอย่างที่จบการศึกษาในระดับปริญญาตรีสายอาชีพศึกษาเพิ่มขึ้นจนถึงปี 2012 แต่กลับลดลงใน 2013 โดยที่ข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 4 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด



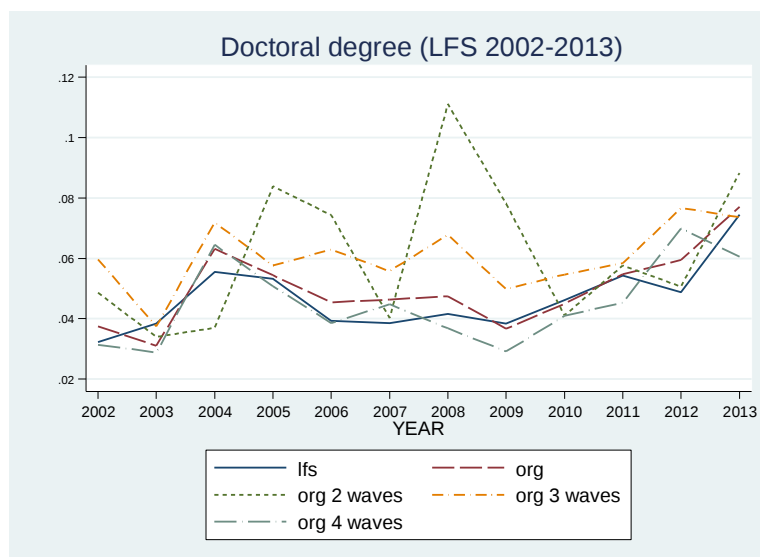
รูปที่ 25: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG
ที่จบการศึกษาปริญญาตรี (สายศึกษาศาสตร์)

จากรูปที่ 25 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่จบการศึกษาระดับปริญญาตรีสายศึกษาศาสตร์พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG ค่อนข้างคงที่ นั่นคือ จำนวนกลุ่มตัวอย่างที่จบการศึกษาในระดับปริญญาตรีสายศึกษาศาสตร์มีค่าคงที่จนถึงปี 2011 แต่มีค่าลดลงในปี 2012-2013 โดยที่ข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 4 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด



รูปที่ 26: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG
ที่จบการศึกษาปริญญาโท

จากรูปที่ 26 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่จบการศึกษาระดับปริญญาโท พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มเพิ่มขึ้นตลอดระยะเวลา นั่นคือ จำนวนกลุ่มตัวอย่างที่จบการศึกษาในระดับปริญญาโทเพิ่มขึ้นและคาดว่าจะเพิ่มขึ้นอีกในปีถัดไป โดยที่ข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 4 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด

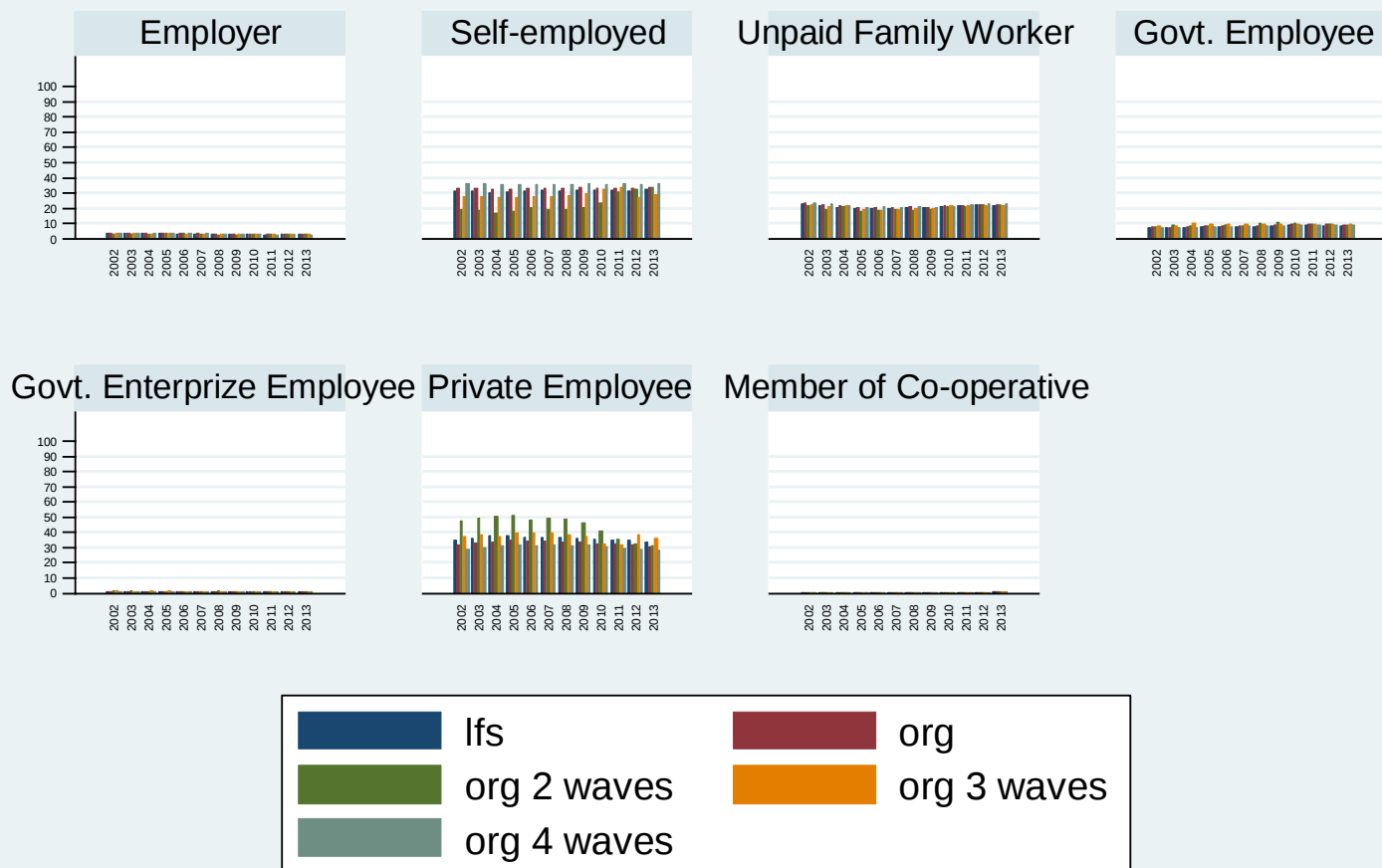


รูปที่ 27: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่จบการศึกษาระดับปริญญาเอก

รูปที่ 27 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่จบการศึกษาระดับปริญญาเอก พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG ค่อยข้างคงที่ในช่วงปี 2002-2011 และเพิ่มขึ้นในปี 2012-2013 นั่นคือ จำนวนกลุ่มตัวอย่างที่จบการศึกษาในระดับปริญญาเอกเพิ่มขึ้นในปี 2012-2013 โดยที่ข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 4 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด

เมื่อพิจารณาร้อยละของจำนวนกลุ่มตัวอย่างที่แบ่งตามกลุ่มอาชีพ จากรูปที่ 28 พบว่า ทั้งข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีคนที่ประกอบอาชีพลูกจ้างเอกชนสูงที่สุด คิดเป็นร้อยละ 34-37 รองลงมาเป็นคนประกอบธุรกิจส่วนตัวโดยไม่มีลูกจ้าง คิดเป็นร้อยละ 29-33 และลำดับที่ 3 คือคนที่ช่วยธุรกิจครัวเรือน โดยไม่ได้รับค่าจ้าง คิดเป็นร้อยละ 20-22 และลำดับที่ 4 คือคนที่ประกอบอาชีพลูกจ้างรัฐบาล คิดเป็นร้อยละ 8-9 และลำดับที่ 5 คือคนที่ประกอบอาชีพนายจ้าง คิดเป็นร้อยละ 3-4 ตามลำดับ

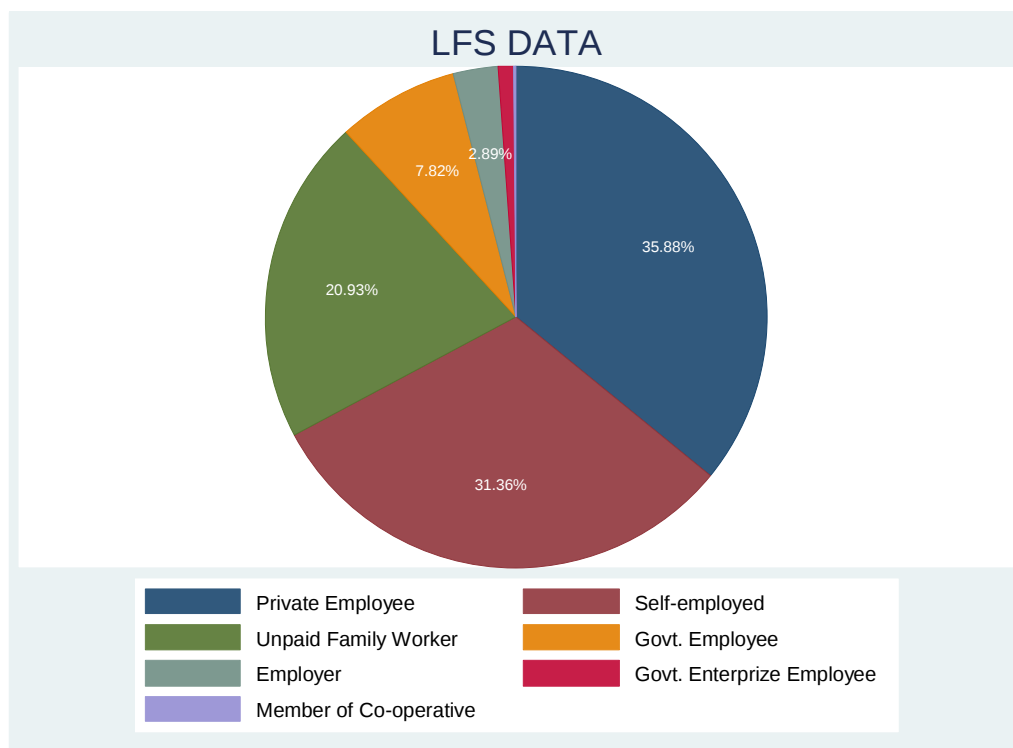
LFS 2002-2013



Graphs by Work Status (2001-recent)

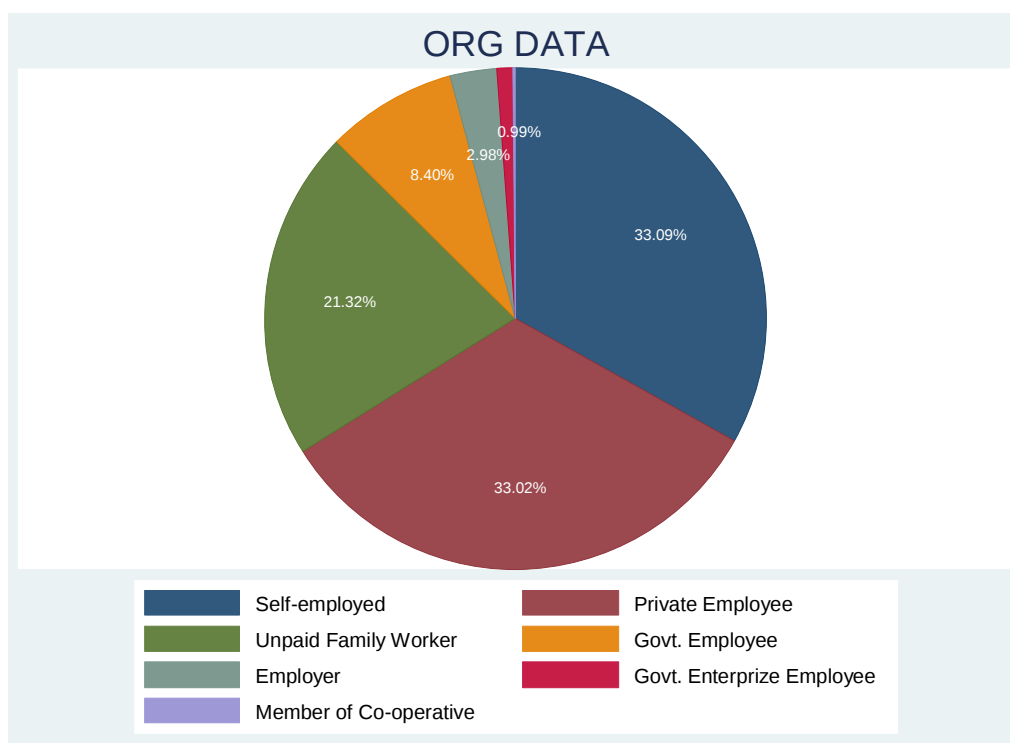
รูปที่ 28: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG แบ่งตามกลุ่มอาชีพ

ร้อยละของคนที่ประกอบอาชีพในกลุ่มต่าง ๆ ที่แยกตามชุดข้อมูล ได้แก่ LFS, ORG, ORG 2 waves, ORG 3 waves และ ORG 4 waves สามารถแสดงได้ในรูปที่ 29 - 33 ดังต่อไปนี้



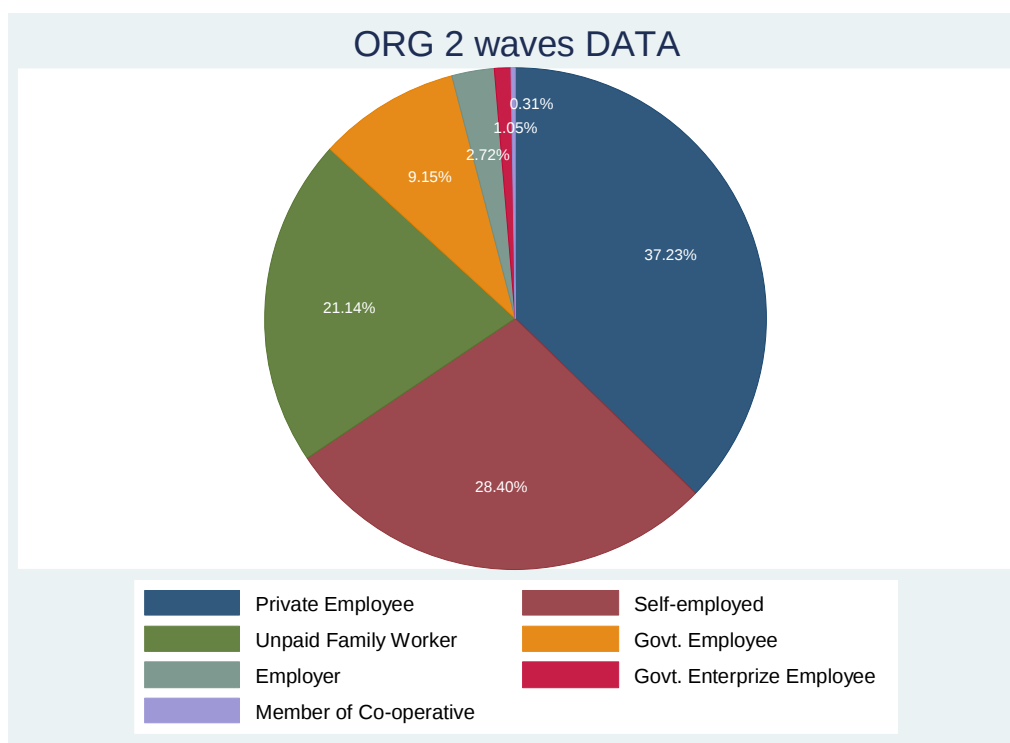
รูปที่ 29: ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS ตั้งแต่ปี 2002-2013
แบ่งตามกลุ่มอาชีพ

จากรูปที่ 29 แสดงร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างที่แบ่งตามกลุ่มอาชีพ จากข้อมูล LFS ตั้งแต่ปี 2002-2013 พบว่า คนที่ประกอบอาชีพลูกจ้างเอกชนมีร้อยละสูงที่สุด คิดเป็นร้อยละ 35.88 รองลงมา เป็นคนที่ประกอบธุรกิจส่วนตัวโดยไม่มีลูกจ้าง คิดเป็นร้อยละ 31.16 และลำดับที่ 3 คือคนที่ช่วยธุรกิจครัวเรือนโดยไม่ได้รับค่าจ้าง คิดเป็นร้อยละ 20.93 และลำดับที่ 4 คือคนที่ประกอบอาชีพลูกจ้างรัฐบาล คิดเป็นร้อยละ 7.82 และลำดับที่ 5 คือคนที่ประกอบอาชีพนายจ้าง คิดเป็นร้อยละ 2.89 และลำดับที่ 6 คือคนที่ประกอบอาชีพลูกจ้างรัฐวิสาหกิจ และลำดับสุดท้ายคือคนที่ประกอบอาชีพโดยการรวมกลุ่ม ตามลำดับ



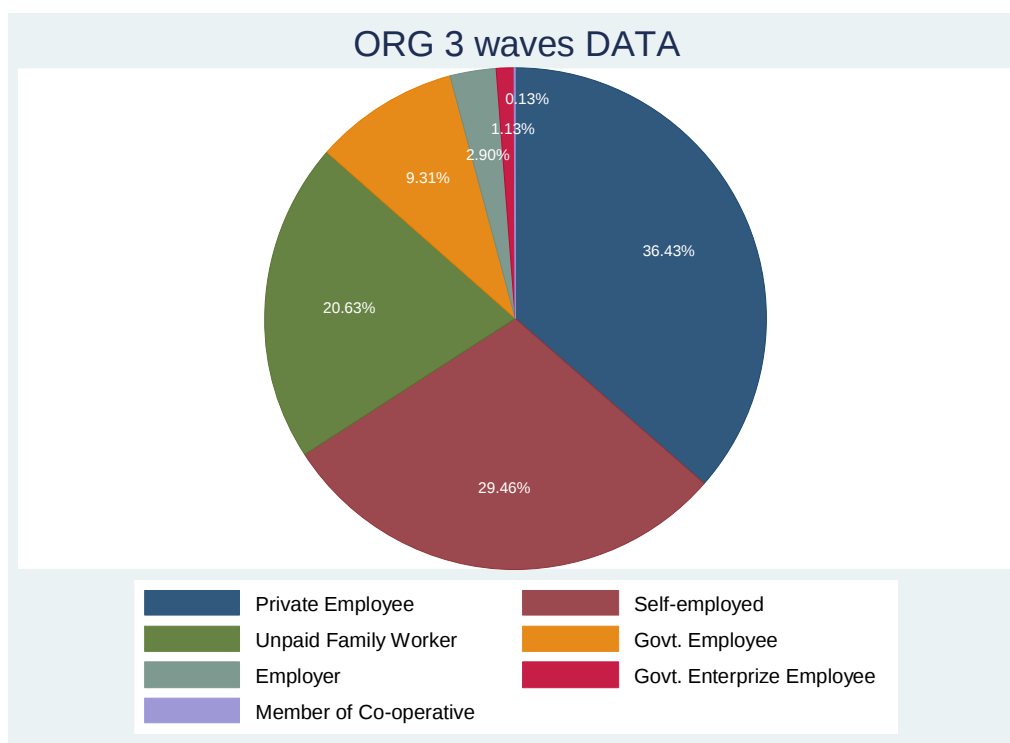
รูปที่ 30: ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG ตั้งแต่ปี 2002-2013
แบ่งตามกลุ่มอาชีพ

จากรูปที่ 30 แสดงร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างที่แบ่งตามกลุ่มอาชีพ จากข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์มากกว่า 1 ครั้งขึ้นไป ตั้งแต่ปี 2002-2013 พบว่า คนที่ประกอบอาชีพลูกจ้างเอกชนมีร้อยละสูงสุด คิดเป็นร้อยละ 33.09 รองลงมาเป็นคนประกอบธุรกิจส่วนตัวโดยไม่มีลูกจ้าง คิดเป็นร้อยละ 33.02 และลำดับที่ 3 คือคนที่ช่วยธุรกิจครัวเรือนโดยไม่ได้รับค่าจ้าง คิดเป็นร้อยละ 21.32 และลำดับที่ 4 คือคนที่ประกอบอาชีพลูกจ้างรัฐบาล คิดเป็นร้อยละ 8.40 และลำดับที่ 5 คือคนที่ประกอบอาชีพนายจ้าง คิดเป็นร้อยละ 2.98 และลำดับที่ 6 คือคนที่ประกอบอาชีพลูกจ้างรัฐวิสาหกิจ และลำดับสุดท้ายคือคนที่ประกอบอาชีพโดยการรวมกลุ่ม ตามลำดับ



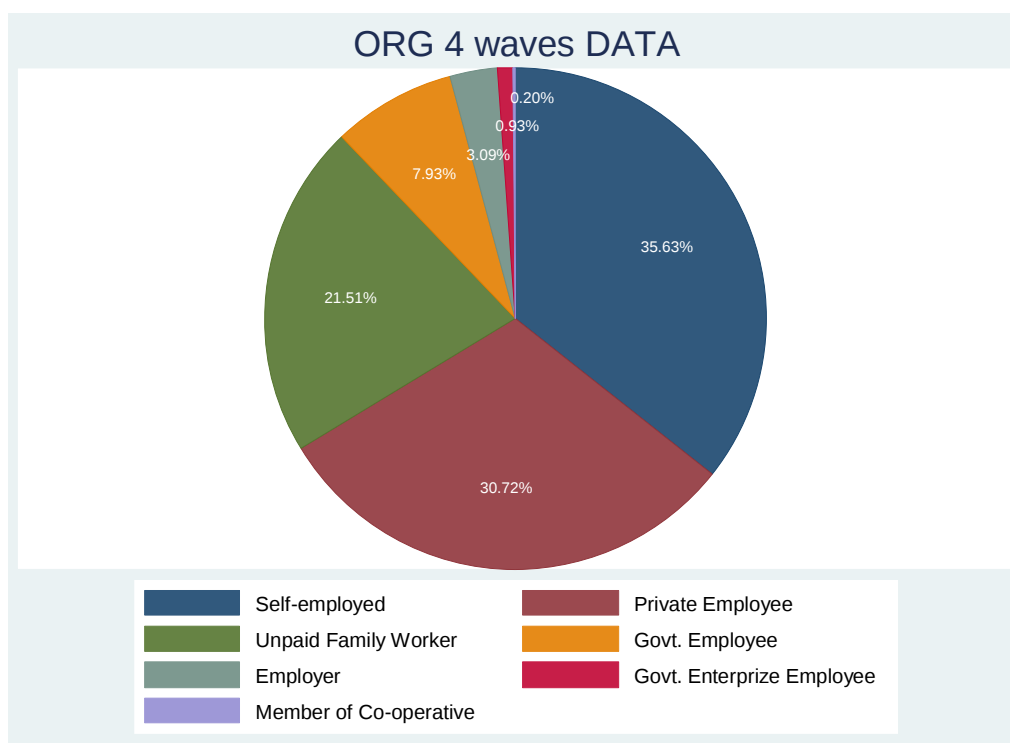
รูปที่ 31: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 2 waves ตั้งแต่ปี 2002-2013
แบ่งตามกลุ่มอาชีพ

จากรูปที่ 31 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่แบ่งตามกลุ่มอาชีพ จากข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์ 2 ครั้ง ตั้งแต่ปี 2002-2013 พบว่า คนที่ประกอบอาชีพลูกจ้างเอกชนมีร้อยละสูงสุด คิดเป็นร้อยละ 37.23 รองลงมาเป็นคนประกอบธุรกิจส่วนตัวโดยไม่มีลูกจ้าง คิดเป็นร้อยละ 28.40 และลำดับที่ 3 คือคนที่ช่วยธุรกิจครัวเรือนโดยไม่ได้รับค่าจ้าง คิดเป็นร้อยละ 21.14 และลำดับที่ 4 คือคนที่ประกอบอาชีพลูกจ้างรัฐบาล คิดเป็นร้อยละ 9.15 และลำดับที่ 5 คือคนที่ประกอบอาชีพนายจ้าง คิดเป็นร้อยละ 2.72 และลำดับที่ 6 คือคนที่ประกอบอาชีพลูกจ้างรัฐวิสาหกิจ และลำดับสุดท้ายคือคนที่ประกอบอาชีพโดยการรวมกลุ่มตามลำดับ



รูปที่ 32: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 3 waves ตั้งแต่ปี 2002-2013
แบ่งตามกลุ่มอาชีพ

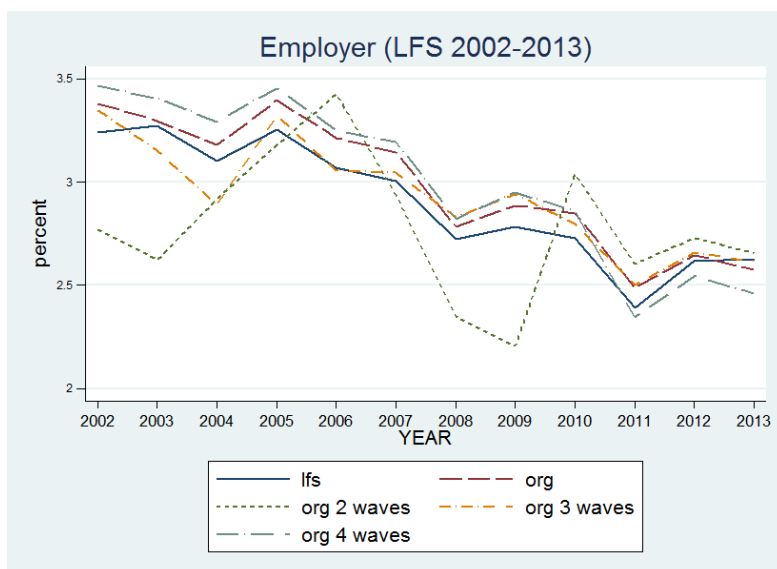
จากรูปที่ 32 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่แบ่งตามกลุ่มอาชีพ จากข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์ 3 ครั้ง ตั้งแต่ปี 2002-2013 พบว่า คนที่ประกอบอาชีพลูกจ้างเอกชนมีร้อยละสูงที่สุด คิดเป็นร้อยละ 36.43 รองลงมาเป็นคนประกอบธุรกิจส่วนตัวโดยไม่มีลูกจ้าง คิดเป็นร้อยละ 29.46 และลำดับที่ 3 คือคนที่ช่วยธุรกิจครัวเรือนโดยไม่ได้รับค่าจ้าง คิดเป็นร้อยละ 20.63 และลำดับที่ 4 คือคนที่ประกอบอาชีพลูกจ้างรัฐบาล คิดเป็นร้อยละ 9.13 และลำดับที่ 5 คือคนที่ประกอบอาชีพนายจ้าง คิดเป็นร้อยละ 2.90 และลำดับที่ 6 คือคนที่ประกอบอาชีพลูกจ้างรัฐวิสาหกิจ และลำดับสุดท้ายคือคนที่ประกอบอาชีพโดยการรวมกลุ่มตามลำดับ



รูปที่ 33: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 4 waves ตั้งแต่ปี 2002-2013
แบ่งตามกลุ่มอาชีพ

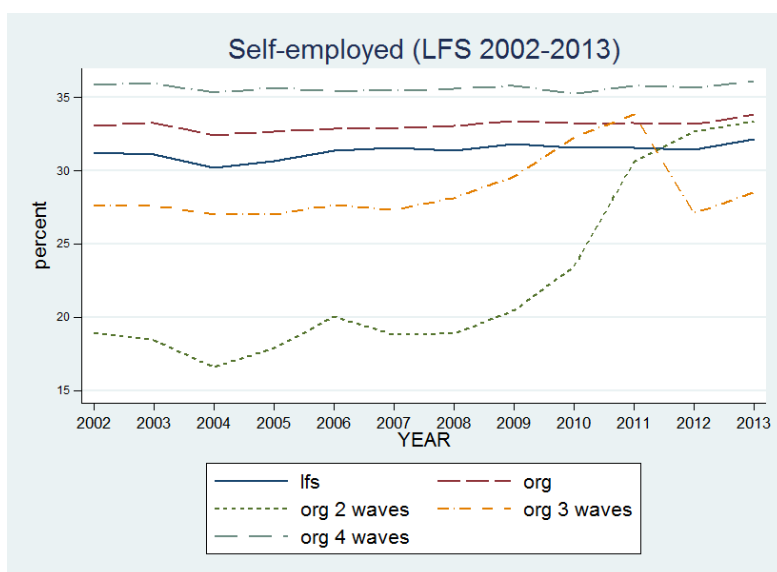
จากรูปที่ 33 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่แบ่งตามกลุ่มอาชีพ จากข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์ 4 ครั้ง ตั้งแต่ปี 2002-2013 พบว่า คนที่ประกอบอาชีพลูกจ้างเอกชนมีร้อยละสูงที่สุด คิดเป็นร้อยละ 35.63 รองลงมาเป็นคนประกอบธุรกิจส่วนตัวโดยไม่มีลูกจ้าง คิดเป็นร้อยละ 30.72 และลำดับที่ 3 คือคนที่ช่วยธุรกิจครัวเรือนโดยไม่ได้รับค่าจ้าง คิดเป็นร้อยละ 21.51 และลำดับที่ 4 คือคนที่ประกอบอาชีพลูกจ้างรัฐบาล คิดเป็นร้อยละ 7.93 และลำดับที่ 5 คือคนที่ประกอบอาชีพนายจ้าง คิดเป็นร้อยละ 3.09 และลำดับที่ 6 คือคนที่ประกอบอาชีพลูกจ้างรัฐวิสาหกิจ และลำดับสุดท้ายคือคนที่ประกอบอาชีพโดยการรวมกลุ่มตามลำดับ

ส่วนถัดไปจะเป็นการเปรียบเทียบกราฟเส้นที่แสดงร้อยละโดยเฉลี่ยต่อปีของแต่ละชุดข้อมูลที่แบ่งตามกลุ่มอาชีพ ซึ่งสามารถแสดงได้ดังต่อไปนี้



รูปที่ 34: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่ประกอบอาชีพนายจ้าง

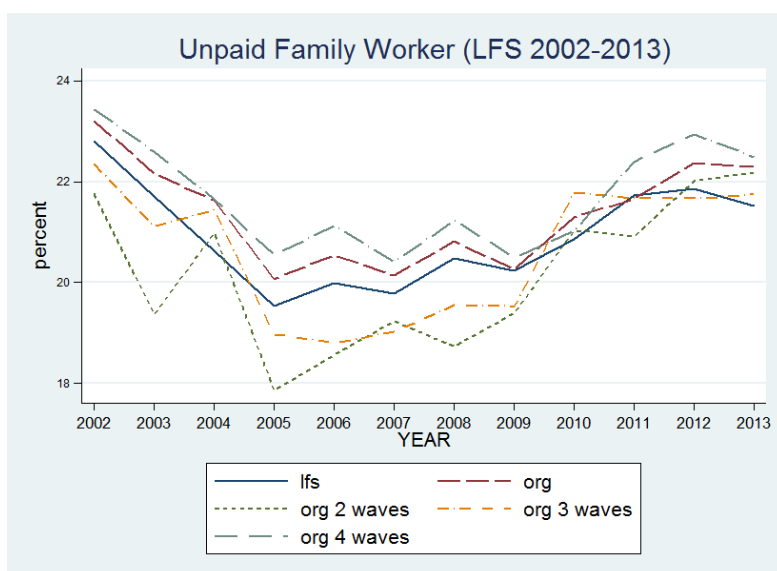
จากรูปที่ 34 แสดงร้อยละของคนที่ประกอบอาชีพนายจ้าง พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มลดลง ซึ่งจะเห็นได้ว่าร้อยละของคนที่ประกอบอาชีพนายจ้างจากข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 3 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด



รูปที่ 35: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่ประกอบธุรกิจส่วนตัวโดยไม่มีลูกจ้าง

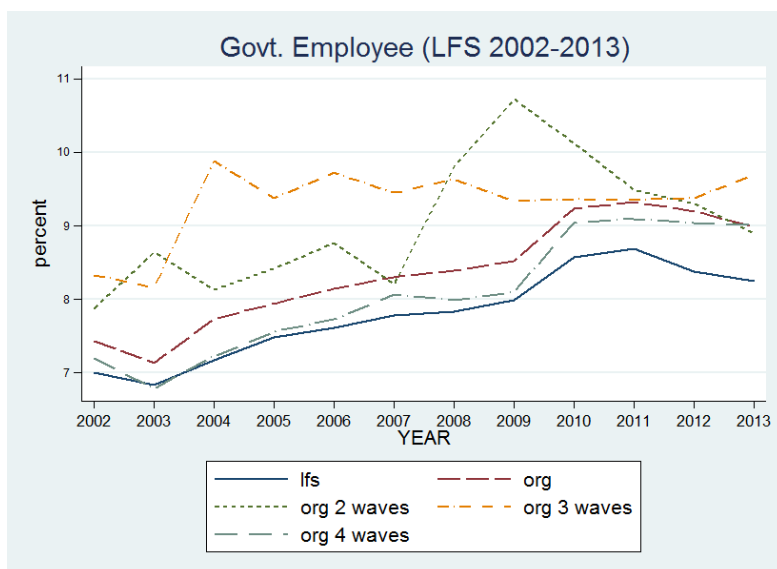
จากรูปที่ 35 แสดงร้อยละของคนที่ประกอบธุรกิจส่วนตัวโดยไม่มีลูกจ้าง พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG, ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 3 ครั้ง และ 4 ครั้ง มีแนวโน้มคงที่ แต่กลับพบว่า ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 2 ครั้ง มีแนวโน้มร้อยละของคนที่ประกอบธุรกิจส่วนตัวโดยไม่มีลูกจ้าง

เพิ่มขึ้น และมีค่าสูงกว่าข้อมูล LFS ในระหว่างปี 2012-2013 และจะเห็นได้ว่าร้อยละของคนประกอบธุรกิจส่วนตัวโดยไม่มีลูกจ้างจากข้อมูล ORG ที่กลุ่มตัวอย่างถูกสัมภาษณ์ 4 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด



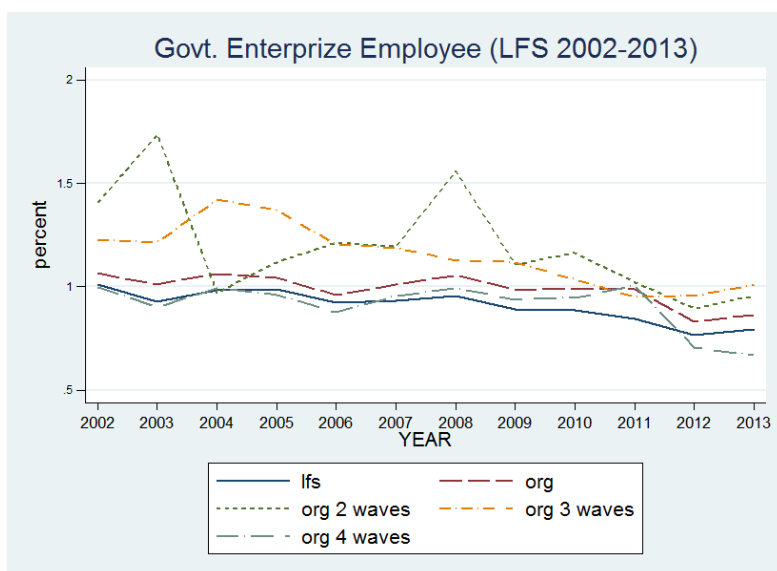
รูปที่ 36: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่ช่วยธุรกิจครัวเรือนโดยไม่ได้รับค่าจ้าง

จากรูปที่ 36 แสดงร้อยละของคนช่วยธุรกิจครัวเรือนโดยไม่ได้รับค่าจ้าง พบว่า ทั้งข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มลดลงในช่วงปี 2002-2005 และมีแนวโน้มเพิ่มสูงขึ้นในช่วงปี 2006-2013 โดยที่ร้อยละของคนช่วยธุรกิจครัวเรือนโดยไม่ได้รับค่าจ้างจากข้อมูลตัวอย่างซ้ำ ORG ที่กลุ่มตัวอย่างถูกสัมภาษณ์มากกว่า 1 ครั้งมีค่าใกล้เคียงกับข้อมูล LFS มากกว่าชุดข้อมูลที่แยกกลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ออกเป็น 2 ครั้ง, 3 ครั้ง และ 4 ครั้ง ตามลำดับ



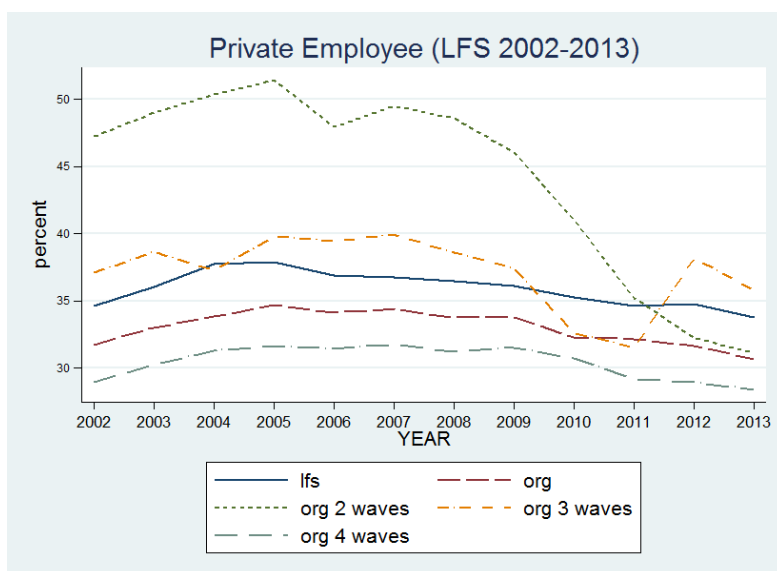
รูปที่ 37: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG
ที่ประกอบอาชีพลูกจ้างรัฐบาล

จากรูปที่ 37 แสดงร้อยละของคนที่ประกอบอาชีพลูกจ้างรัฐบาล พบว่า ข้อมูล LFS และข้อมูล ตัวอย่างซ้ำ ORG มีแนวโน้มเพิ่มขึ้น ซึ่งจะเห็นได้ว่าร้อยละของคนที่ประกอบอาชีพประกอบอาชีพลูกจ้าง รัฐบาลจากข้อมูล ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 4 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด



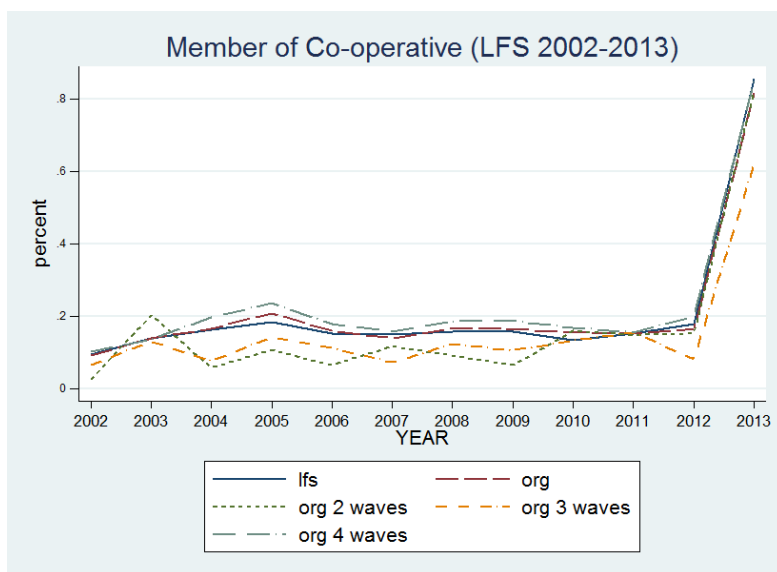
รูปที่ 38: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG
ที่ประกอบอาชีพลูกจ้างรัฐวิสาหกิจ

จากรูปที่ 38 แสดงร้อยละของคนที่ประกอบอาชีพลูกจ้างรัฐวิสาหกิจ พบว่า ข้อมูล LFS และข้อมูล ตัวอย่างซ้ำ ORG มีแนวโน้มลดลง ซึ่งจะเห็นได้ว่าร้อยละของคนที่ประกอบอาชีพลูกจ้างรัฐวิสาหกิจจากข้อมูล ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 4 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด ในขณะที่ร้อยละของคนที่ประกอบอาชีพนายจ้างจากข้อมูล ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 2 ครั้ง มีค่าสูงที่สุด



รูปที่ 39: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG
ที่ประกอบอาชีพลูกจ้างเอกชน

จากรูปที่ 39 แสดงร้อยละของคนที่ประกอบอาชีพลูกจ้างเอกชน พบว่า ข้อมูล LFS และข้อมูล ตัวอย่างซ้ำ ORG มีแนวโน้มคงที่ ซึ่งจะเห็นได้ว่าร้อยละของคนที่ประกอบอาชีพลูกจ้างเอกชนจากข้อมูล ORG ที่กลุ่มตัวอย่างถูกสัมภาษณ์มากกว่า 1 ครั้งมีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด ในขณะที่ร้อยละของคนที่ประกอบอาชีพลูกจ้างเอกชนจากข้อมูล ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 2 ครั้ง มีค่าสูงที่สุด และ ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 4 ครั้ง มีค่าต่ำที่สุด

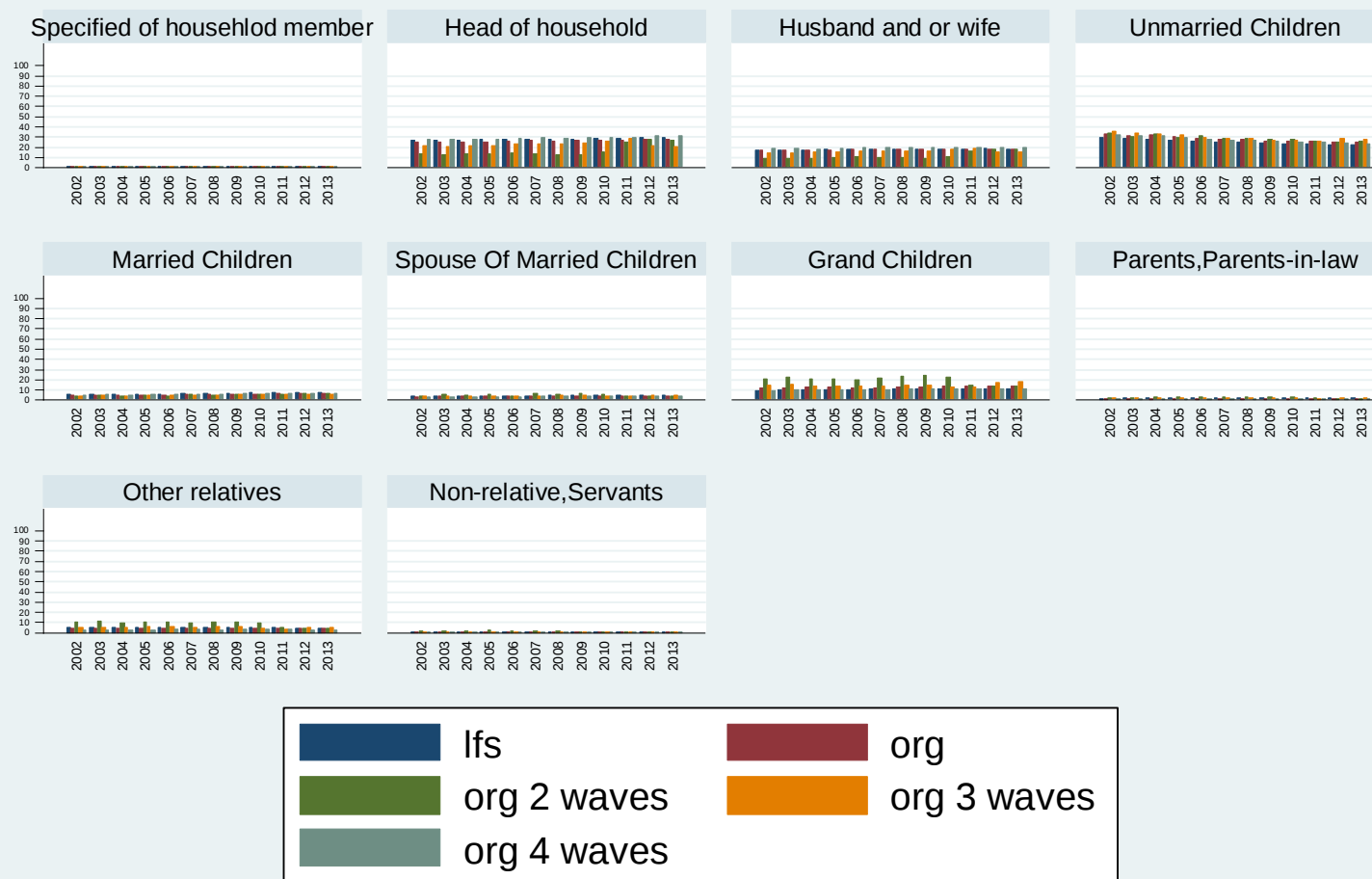


รูปที่ 40: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG
ที่ประกอบอาชีพโดยการรวมกลุ่ม

จากรูปที่ 40 แสดงร้อยละของคนที่ประกอบอาชีพโดยการรวมกลุ่ม พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มคงที่ ในช่วงปี 2002-2013 แต่เพิ่มสูงขึ้นในปี 2013 ซึ่งจะเห็นได้ว่าร้อยละของคนที่ประกอบอาชีพโดยการรวมกลุ่มจากทั้งข้อมูล ORG ที่กลุ่มตัวอย่างถูกสัมภาษณ์มากกว่า 1 ครั้ง, 2 ครั้ง, 3 ครั้ง และ 4 ครั้งมีค่าใกล้เคียงกับร้อยละของคนที่ประกอบอาชีพโดยการรวมกลุ่มจากข้อมูล LFS

เมื่อพิจารณาร้อยละของจำนวนกลุ่มตัวอย่างที่แบ่งตามความสัมพันธ์กับหัวหน้าครัวเรือน จากรูปที่ 10 พบว่า ทั้งข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีร้อยละของคนที่เป็หัวหน้าครัวเรือนมากที่สุด คิดเป็นร้อยละ 25-33 รองลงมาเป็นบุตรที่ยังไม่สมรส คิดเป็นร้อยละ 25-28 และลำดับที่ 3 คือคนที่เป็สามีหรือภรรยาของหัวหน้าครัวเรือน คิดเป็นร้อยละ 17-21 และลำดับที่ 4 คือคนที่เป็หลานของหัวหน้าครัวเรือน คิดเป็นร้อยละ 7-12 และลำดับที่ 5 คือคนที่เป็บุตรซึ่งสมรสแล้ว คิดเป็นร้อยละ 5-7 และลำดับที่ 6 คือญาติอื่น ๆ ของหัวหน้าครัวเรือน คิดเป็นร้อยละ 2-4 และลำดับที่ 7 คือคนที่เป็คู่สมรสของบุตรที่แต่งงานแล้ว คิดเป็นร้อยละ 3-5 และลำดับที่ 8 คือ เป็พ่อ หรือแม่ หรือพ่อตา หรือแม่ยาย คิดเป็นร้อยละ 1-2 และลำดับที่ 9 คือผู้อาศัยหรือคนใช้ คิดเป็นร้อยละ 0.01-1.4 ตามลำดับ

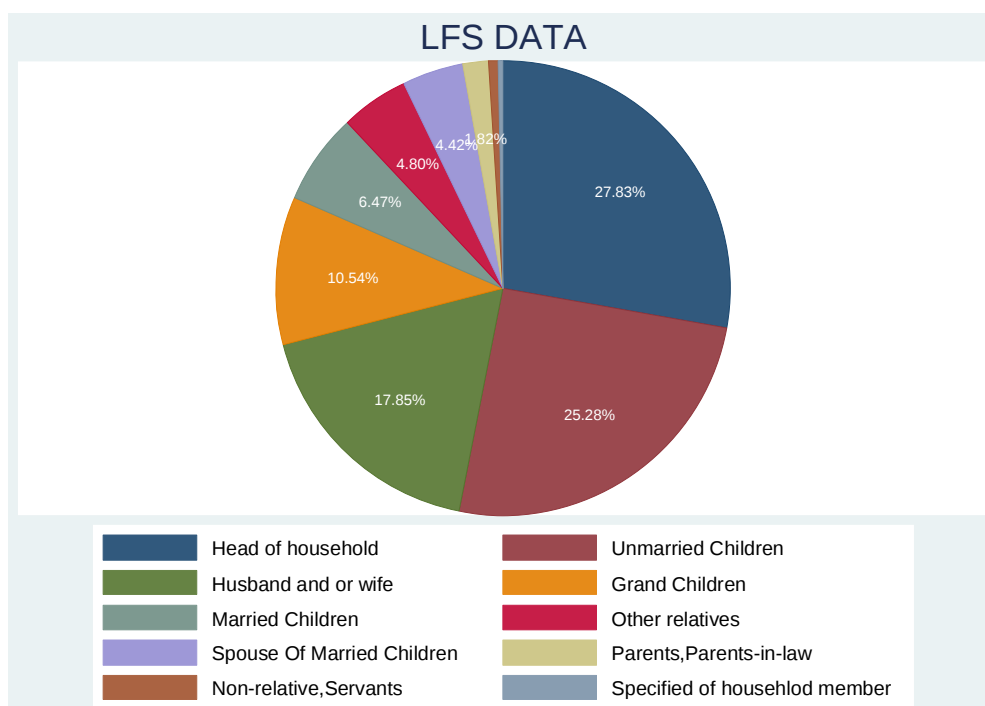
LFS 2002-2013



Graphs by RELATIONSHIP TO HEAD OF HOUSEHOLD (1985-recent)

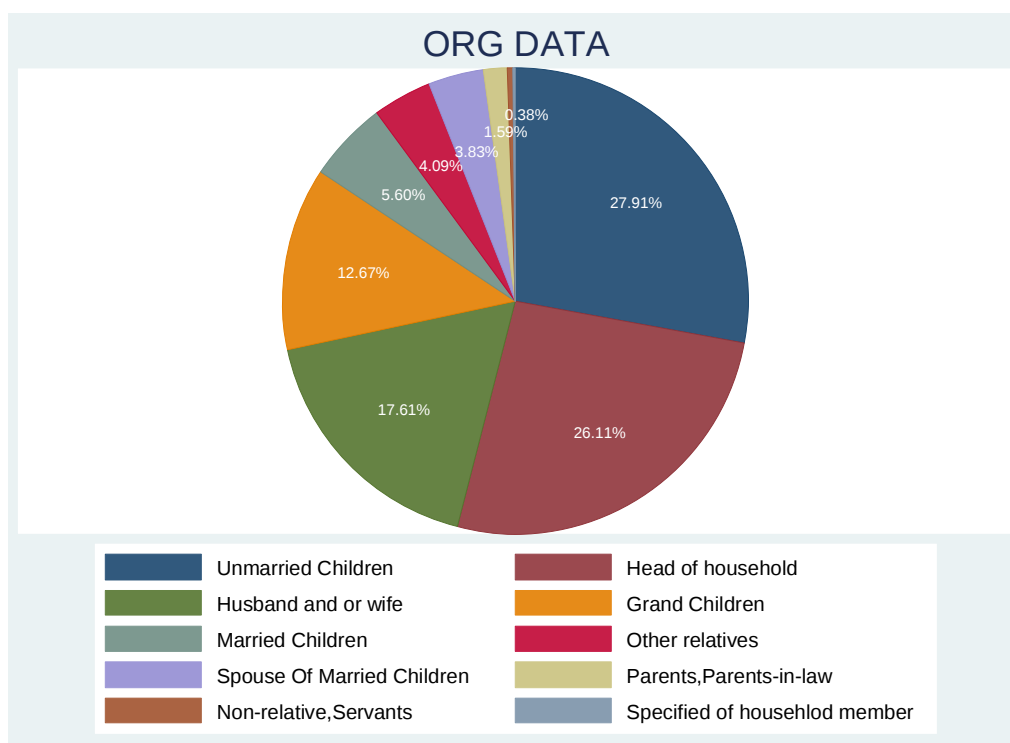
รูปที่ 41: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG
แบ่งตามความสัมพันธ์กับหัวหน้าครัวเรือน

โดยร้อยละของคนที่ประกอบอาชีพในกลุ่มต่าง ๆ แยกตามชุดข้อมูล ได้แก่ LFS, ORG, ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 2 ครั้ง, 3 ครั้ง และ 4 ครั้ง ตามลำดับ สามารถแสดงได้ในรูปที่ 42 - 46 ดังต่อไปนี้



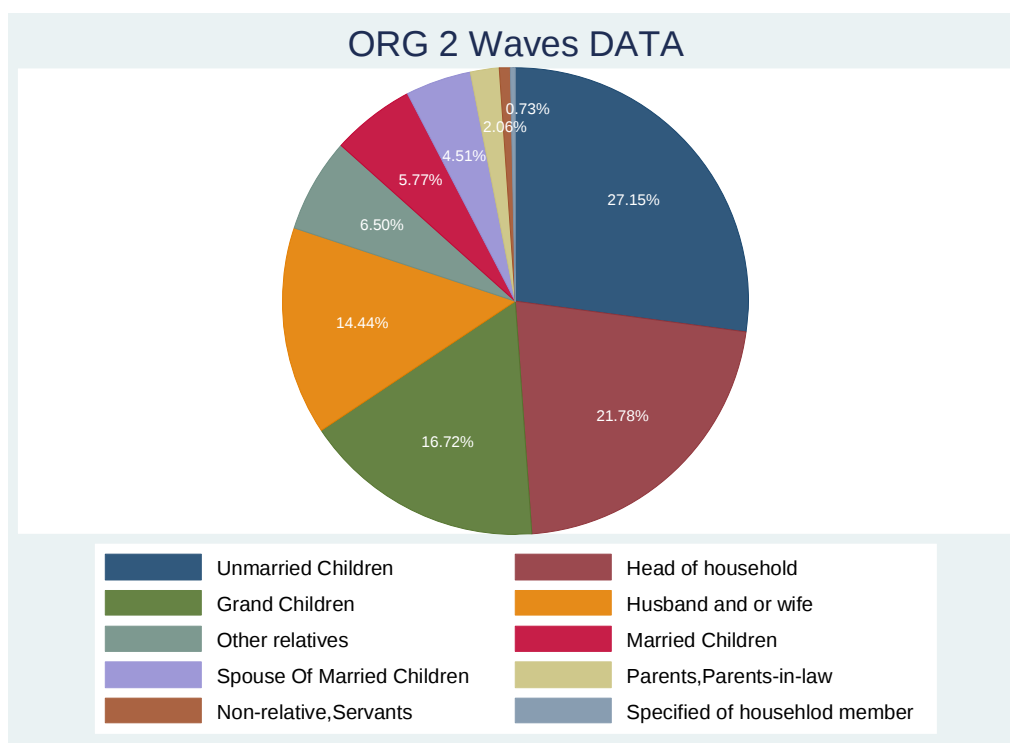
รูปที่ 42: ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS ตั้งแต่ปี 2002-2013
แบ่งตามความสัมพันธ์กับหัวหน้าครัวเรือน

จากรูปที่ 42 แสดงร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS ที่แบ่งตามความสัมพันธ์กับหัวหน้าครัวเรือน พบว่า มีร้อยละของคนที่เป็หัวหน้าครัวเรือนมากที่สุด คิดเป็นร้อยละ 27.83 รองลงมา เป็นบุตรที่ยังไม่สมรส คิดเป็นร้อยละ 25.28 และลำดับที่ 3 คือคนที่เป็สามีหรือภรรยาของหัวหน้าครัวเรือน คิดเป็นร้อยละ 17.85 และลำดับที่ 4 คือคนที่เป็หลานของหัวหน้าครัวเรือน คิดเป็นร้อยละ 10.54 และลำดับที่ 5 คือคนที่เป็บุตรซึ่งสมรสแล้ว คิดเป็นร้อยละ 6.47 และลำดับที่ 6 คือญาติอื่น ๆ ของหัวหน้าครัวเรือน คิดเป็นร้อยละ 4.80 และลำดับที่ 7 คือคนที่เป็คู่สมรสของบุตรที่แต่งงานแล้ว คิดเป็นร้อยละ 4.42 และลำดับที่ 8 คือ เป็พ่อ หรือแม่ หรือพ่อตา หรือแม่ยาย คิดเป็นร้อยละ 1.82 และลำดับที่ 9 คือผู้อาศัยหรือคนใช้ คิดเป็นร้อยละ 0.60 ตามลำดับ



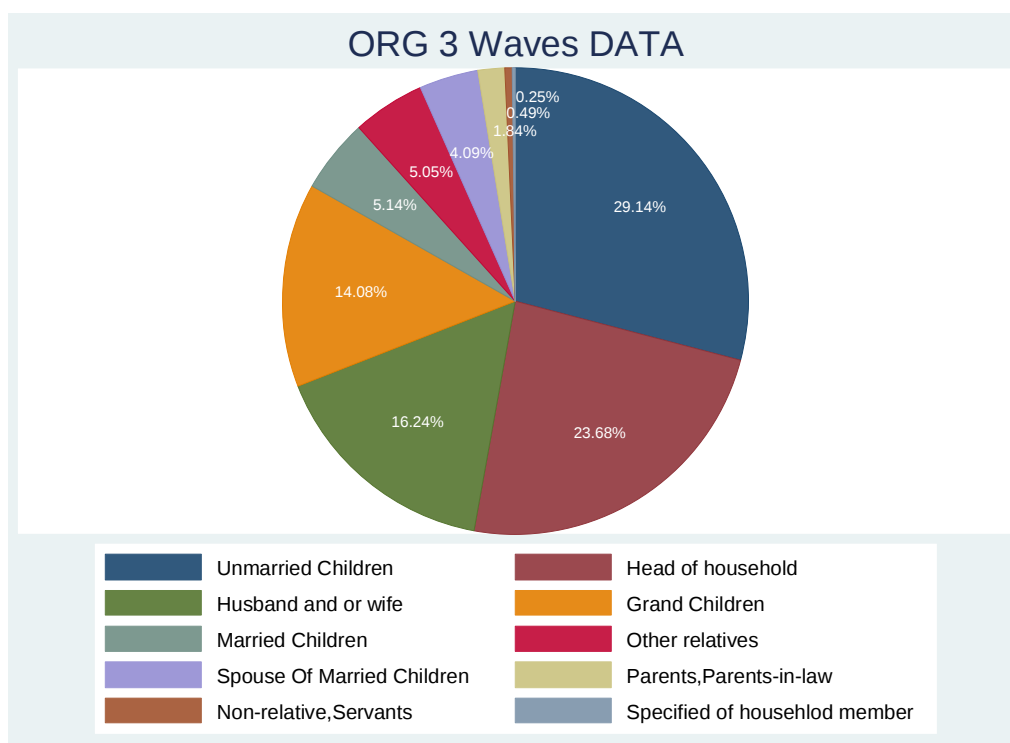
รูปที่ 43: ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG ตั้งแต่ปี 2002-2013
แบ่งตามความสัมพันธ์กับหัวหน้าครัวเรือน

จากรูปที่ 43 แสดงร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์มากกว่า 1 ครั้งขึ้นไป ที่แบ่งตามความสัมพันธ์กับหัวหน้าครัวเรือน พบว่า มีร้อยละของคนที่บุตรที่ยังไม่สมรสมากที่สุด คิดเป็นร้อยละ 27.91 รองลงมาเป็นหัวหน้าครัวเรือนมาก คิดเป็นร้อยละ 26.11 และลำดับที่ 3 คือคนที่เป็นสามีหรือภรรยาของหัวหน้าครัวเรือน คิดเป็นร้อยละ 17.61 และลำดับที่ 4 คือคนที่เป็นหลานของหัวหน้าครัวเรือน คิดเป็นร้อยละ 12.67 และลำดับที่ 5 คือคนที่เป็นบุตรซึ่งสมรสแล้ว คิดเป็นร้อยละ 5.60 และลำดับที่ 6 คือญาติอื่น ๆ ของหัวหน้าครัวเรือน คิดเป็นร้อยละ 4.09 และลำดับที่ 7 คือคนที่เป็นคู่สมรสของบุตรที่แต่งงานแล้ว คิดเป็นร้อยละ 3.83 และลำดับที่ 8 คือ เป็นพ่อ หรือแม่ หรือพ่อตา หรือแม่ยาย คิดเป็นร้อยละ 1.59 และลำดับที่ 9 คือผู้อาศัยหรือคนใช้ คิดเป็นร้อยละ 0.38 ตามลำดับ



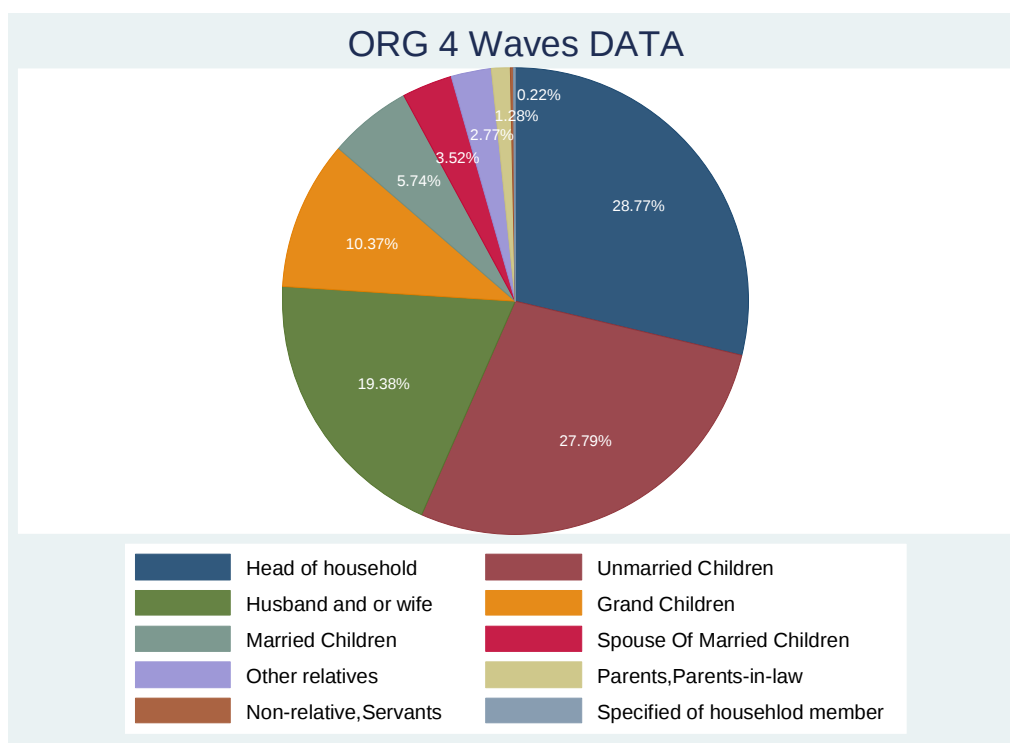
รูปที่ 44: ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 2 waves ตั้งแต่ปี 2002-2013 แบ่งตามความสัมพันธ์กับหัวหน้าครัวเรือน

จากรูปที่ 44 แสดงร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์ 2 ครั้งที่แบ่งตามความสัมพันธ์กับหัวหน้าครัวเรือน พบว่า มีร้อยละของคนที่เป็บุตรที่ยังไม่สมรสมากที่สุด คิดเป็นร้อยละ 27.15 รองลงมาเป็หัวหน้าครัวเรือน คิดเป็นร้อยละ 21.78 และลำดับที่ 3 คือคนที่เป็นหลานของหัวหน้าครัวเรือน คิดเป็นร้อยละ 16.72 และลำดับที่ 4 คือคนที่เป็นสามีหรือภรรยาของหัวหน้าครัวเรือน คิดเป็นร้อยละ 14.44 คนที่ และลำดับที่ 5 คือญาติอื่น ๆ ของหัวหน้าครัวเรือน คิดเป็นร้อยละ 6.50 และลำดับที่ 6 คือคนที่เป็นบุตรซึ่งสมรสแล้ว คิดเป็นร้อยละ 5.77 คือ และลำดับที่ 7 คือคนที่เป็นคู่สมรสของบุตรที่แต่งงานแล้ว คิดเป็นร้อยละ 4.51 และลำดับที่ 8 คือ เป็นพ่อ หรือแม่ หรือพ่อตา หรือแม่ยาย คิดเป็นร้อยละ 2.06 และลำดับที่ 9 คือผู้อาศัยหรือคนใช้ คิดเป็นร้อยละ 0.73 ตามลำดับ



รูปที่ 45: ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 3 waves ตั้งแต่ปี 2002-2013 แบ่งตามความสัมพันธ์กับหัวหน้าครัวเรือน

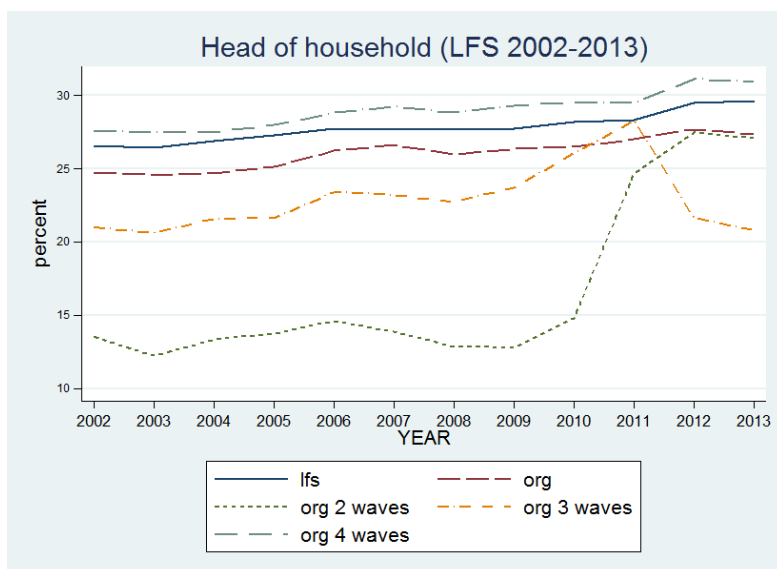
จากรูปที่ 45 แสดงร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์ 3 ครั้งซึ่งแบ่งตามความสัมพันธ์กับหัวหน้าครัวเรือน พบว่า มีร้อยละของคนที่เป็นบุตรที่ยังไม่สมรสมากที่สุด คิดเป็นร้อยละ 29.14 รองลงมาเป็นหัวหน้าครัวเรือน คิดเป็นร้อยละ 23.68 และลำดับที่ 3 คือคนที่ เป็นสามีหรือภรรยาของหัวหน้าครัวเรือน คิดเป็นร้อยละ 16.24 และลำดับที่ 4 คือคนที่ เป็นหลานของหัวหน้าครัวเรือน คิดเป็นร้อยละ 14.08 และลำดับที่ 5 คือคนที่ เป็นบุตรซึ่งสมรสแล้ว คิดเป็นร้อยละ 5.14 และลำดับที่ 6 คือญาติอื่น ๆ ของหัวหน้าครัวเรือน คิดเป็นร้อยละ 5.05 และลำดับที่ 7 คือคนที่ เป็นคู่สมรสของบุตรที่แต่งงานแล้ว คิดเป็นร้อยละ 4.09 และลำดับที่ 8 คือ เป็นพ่อ หรือแม่ หรือพ่อตา หรือแม่ยาย คิดเป็นร้อยละ 1.84 และลำดับที่ 9 คือผู้อาศัยหรือคนใช้ คิดเป็นร้อยละ 0.49 ตามลำดับ



รูปที่ 46: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 4 waves ตั้งแต่ปี 2002-2013
แบ่งตามความสัมพันธ์กับหัวหน้าครัวเรือน

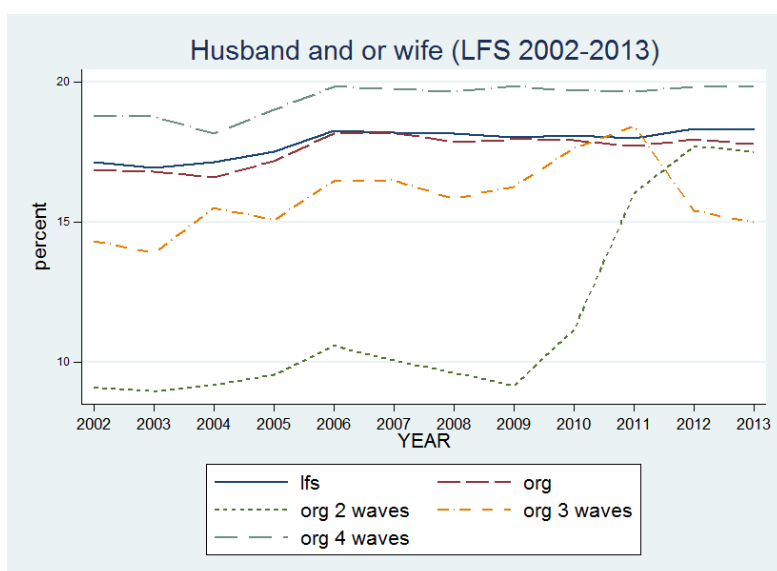
จากรูปที่ 46 แสดงร้อยละของจำนวนกลุ่มตัวอย่างข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์ 4 ครั้ง que แบ่งตามความสัมพันธ์กับหัวหน้าครัวเรือน พบว่า มีร้อยละของคนที่เป็นหัวหน้าครัวเรือนมากที่สุด คิดเป็นร้อยละ 28.77 รองลงมาเป็นบุตรที่ยังไม่สมรส คิดเป็นร้อยละ 27.79 และลำดับที่ 3 คือคนที่เป็นสามีหรือภรรยาของหัวหน้าครัวเรือน คิดเป็นร้อยละ 19.38 และลำดับที่ 4 คือคนที่เป็นหลานของหัวหน้าครัวเรือน คิดเป็นร้อยละ 10.37 และลำดับที่ 5 คือคนที่เป็นบุตรซึ่งสมรสแล้ว คิดเป็นร้อยละ 5.74 และลำดับที่ 6 คือคนที่เป็นคู่สมรสของบุตรที่แต่งงานแล้ว คิดเป็นร้อยละ 3.52 และลำดับที่ 7 คือญาติอื่น ๆ ของหัวหน้าครัวเรือน คิดเป็นร้อยละ 2.77 และลำดับที่ 8 คือ เป็นพ่อ หรือแม่ หรือพ่อตา หรือแม่ยาย คิดเป็นร้อยละ 1.28 และลำดับที่ 9 คือผู้อาศัยหรือคนใช้ คิดเป็นร้อยละ 0.22 ตามลำดับ

ส่วนถัดไปจะเป็นการเปรียบเทียบกราฟเส้นที่แสดงร้อยละโดยเฉลี่ยต่อปีของแต่ละชุดข้อมูลที่แบ่งตามความสัมพันธ์กับหัวหน้าครัวเรือน ซึ่งสามารถแสดงได้ดังต่อไปนี้



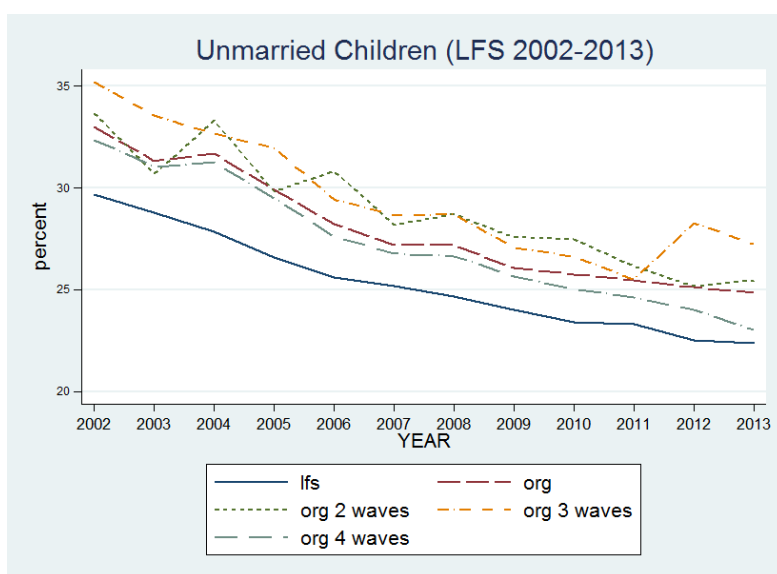
รูปที่ 47: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่เป็นหัวหน้าครัวเรือน

จากรูปที่ 47 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่เป็นหัวหน้าครัวเรือน พบว่า ข้อมูล LFS และ ข้อมูลตัวอย่างซ้ำ ORG, ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 3 ครั้ง และ 4 ครั้ง มีแนวโน้มค่อนข้างคงที่ โดยเฉลี่ยมีคนที่เป็นหัวหน้าครัวเรือนร้อยละ 20-35 ยกเว้นข้อมูลจาก ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 2 ครั้ง ซึ่งร้อยละเฉลี่ยมีแนวโน้มเพิ่มสูงขึ้นในปี 2011-2013 และมีคนที่เป็นหัวหน้าครัวเรือนเฉลี่ยร้อยละ 13-27 ซึ่งมีร้อยละเฉลี่ยน้อยกว่าข้อมูลจากชุดอื่น ๆ



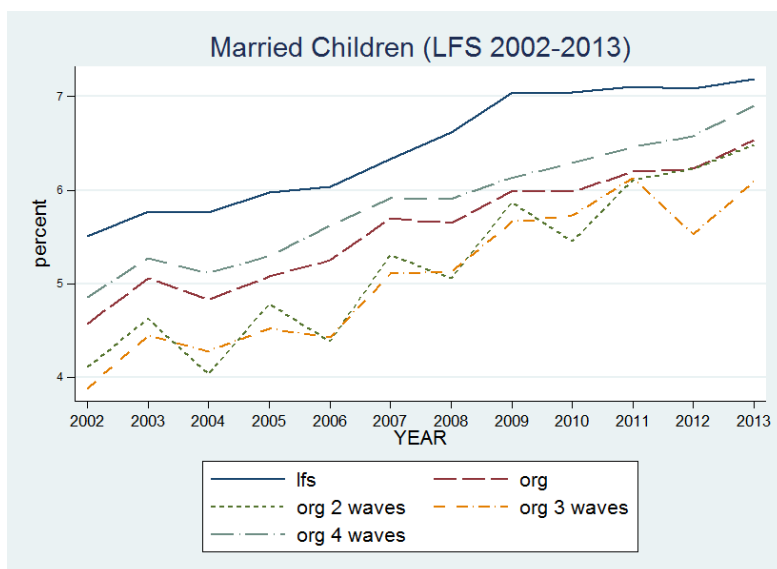
รูปที่ 48: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่เป็นสามีหรือภรรยา

จากรูปที่ 48 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่เป็นสามีหรือภรรยาของหัวหน้าครัวเรือน พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG, ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 3 ครั้ง และ 4 ครั้ง มีแนวโน้มค่อนข้างคงที่ โดยเฉลี่ยมีคนที่เป็นสามีหรือภรรยาของหัวหน้าครัวเรือนร้อยละ 14-20 ยกเว้นข้อมูลจาก ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 2 ครั้ง โดยพบว่า ร้อยละโดยเฉลี่ยมีแนวโน้มเพิ่มสูงขึ้นในปี 2011-2013 ซึ่งมีคนที่เป็นสามีและภรรยาของหัวหน้าครัวเรือนเฉลี่ยร้อยละ 8-17 ซึ่งมีร้อยละเฉลี่ยน้อยกว่าข้อมูลจากชุดอื่น ๆ นอกจากนี้ยังพบว่า ข้อมูล ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 2 ครั้ง มีร้อยละของคนที่เป็นหัวหน้าครัวเรือน และคนที่เป็นสามีหรือภรรยาของหัวหน้าครัวเรือนแตกต่างจากข้อมูล LFS อย่างชัดเจน



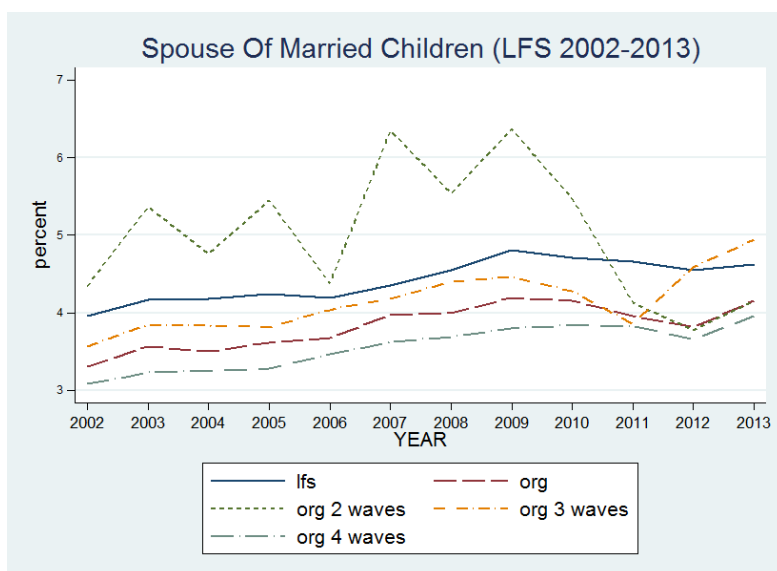
รูปที่ 49: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่เป็นบุตรที่ยังไม่สมรส

จากรูปที่ 49 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่เป็นบุตรที่ยังไม่สมรส พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG, ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 2 ครั้ง, 3 ครั้ง และ 4 ครั้ง มีแนวโน้มลดลงอย่างต่อเนื่อง โดยมีคนที่เป็นบุตรที่ยังไม่สมรสเฉลี่ยร้อยละ 22-35 ซึ่งพบว่าข้อมูลตัวอย่างซ้ำที่มีร้อยละโดยเฉลี่ยของบุตรที่ยังไม่สมรสใกล้เคียงกับข้อมูล LFS มากที่สุดคือ ข้อมูล ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 4 ครั้ง



รูปที่ 50: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่เป็นบุตรที่สมรสแล้ว

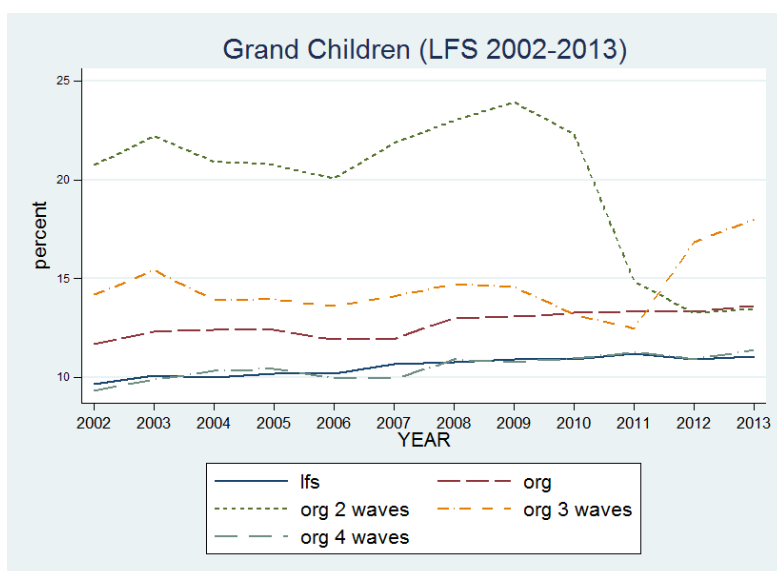
จากรูปที่ 50 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่เป็นบุตรที่สมรสแล้ว พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG, ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 2 ครั้ง, 3 ครั้ง และ 4 ครั้ง มีแนวโน้มเพิ่มขึ้นอย่างต่อเนื่อง โดยมีคนที่เป็นบุตรที่สมรสแล้วเฉลี่ยร้อยละ 3-8 ซึ่งพบว่าข้อมูลตัวอย่างซ้ำที่มีร้อยละโดยเฉลี่ยของบุตรที่สมรสแล้วใกล้เคียงกับข้อมูล LFS มากที่สุดคือ ข้อมูล ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 4 ครั้ง



รูปที่ 51: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่เป็นบุตรเขยหรือบุตรสะใภ้

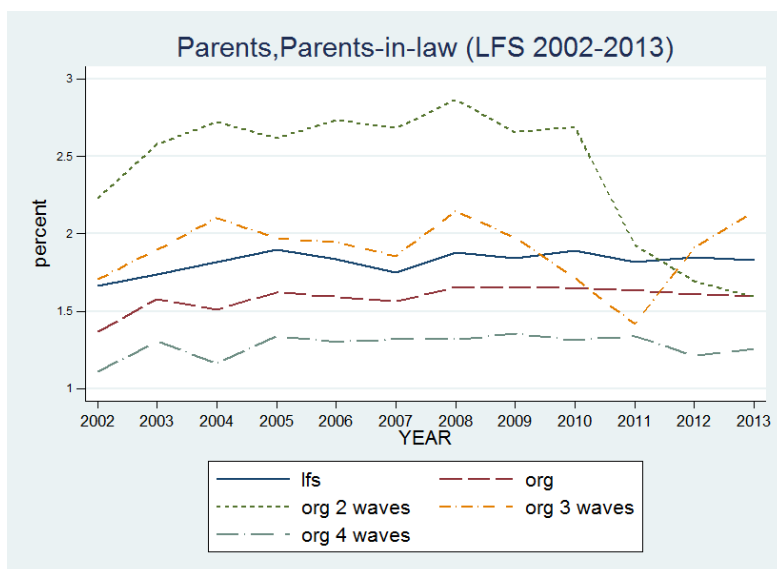
จากรูปที่ 51 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่เป็นบุตรเขยหรือบุตรสะใภ้พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG, ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 3 ครั้ง และ 4 ครั้ง มีแนวโน้มคงที่ โดยมีคนที่เป็นบุตรเขยหรือบุตรสะใภ้เฉลี่ยร้อยละ 3-5 ยกเว้นข้อมูลตัวอย่างซ้ำ ORG 2 waves ที่มีร้อยละเฉลี่ยของคนที่เป็น

บุตรเขยหรือบุตรสะใภ้สูงกว่าข้อมูลชุดอื่น ๆ ซึ่งมีค่าเฉลี่ยร้อยละ 4-6 ซึ่งจากการเปรียบเทียบข้อมูลตัวอย่างซ้ำ ORG กับข้อมูล LFS พบว่าข้อมูลตัวอย่างซ้ำที่มีร้อยละโดยเฉลี่ยของบุตรเขยหรือบุตรสะใภ้ใกล้เคียงกับข้อมูล LFS มากที่สุดคือ ข้อมูล ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 3 ครั้ง



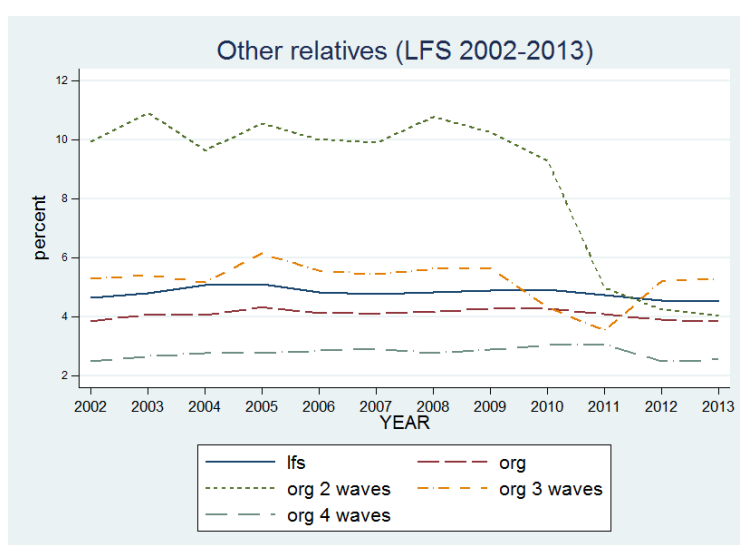
รูปที่ 52: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่เป็นบุตรของบุตร

จากรูปที่ 52 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่เป็นบุตรของบุตร พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG, ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 3 ครั้ง และ 4 ครั้ง มีแนวโน้มคงที่ โดยมีคนที่เป็นบุตรของบุตรเฉลี่ยร้อยละ 9-17 ยกเว้นข้อมูลตัวอย่างซ้ำ ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 2 ครั้ง ที่มีร้อยละเฉลี่ยของคนที่เป็นบุตรของบุตรสูงกว่าข้อมูลชุดอื่น ๆ ซึ่งมีค่าเฉลี่ยร้อยละ 13-24 ซึ่งจากการเปรียบเทียบข้อมูลตัวอย่างซ้ำ ORG กับข้อมูล LFS พบว่าข้อมูลตัวอย่างซ้ำที่มีร้อยละโดยเฉลี่ยของบุตรของบุตรใกล้เคียงกับข้อมูล LFS มากที่สุดคือ ข้อมูล ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 4 ครั้ง



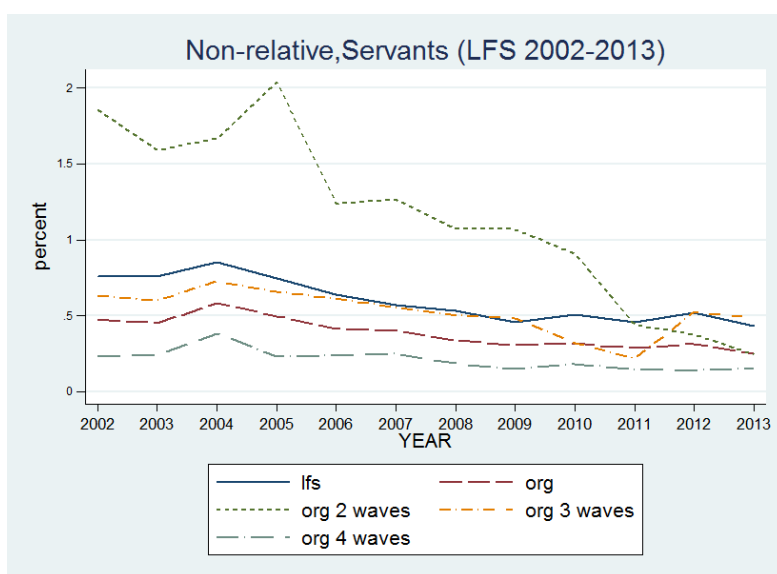
รูปที่ 53: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่เป็นพ่อ หรือแม่ หรือพ่อตา หรือแม่ยาย

จากรูปที่ 53 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่เป็นพ่อ หรือแม่ หรือพ่อตา หรือแม่ยาย พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มคงที่ ซึ่งแต่ละชุดข้อมูลมีร้อยละเฉลี่ยค่อนข้างแตกต่างกัน อย่างชัดเจน โดยจะเห็นได้ว่าร้อยละเฉลี่ยของคนที่เป็นพ่อ หรือแม่ หรือพ่อตา หรือแม่ยายจากข้อมูล ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 2 ครั้ง มีค่าสูงที่สุดอยู่ที่ร้อยละ 1.6-2.8 ในขณะที่ข้อมูล LFS มีร้อยละเฉลี่ยค่อนข้าง คงที่อยู่ที่ร้อยละ 1.5-1.7 และข้อมูล ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 3 ครั้ง มีค่าร้อยละเฉลี่ยอยู่ที่ 1.7-2 และ ข้อมูล ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 4 ครั้ง มีร้อยละเฉลี่ยอยู่ที่ 1.1-1.3 ตามลำดับ โดยที่ข้อมูล ORG กลุ่ม ตัวอย่างถูกสัมภาษณ์ 3 ครั้ง จะมีร้อยละเฉลี่ยของคนที่เป็นพ่อ หรือแม่ หรือพ่อตา หรือแม่ยาย ใกล้เคียงกับ ข้อมูล LFS มากที่สุด



รูปที่ 54: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่เป็นญาติอื่น ๆ

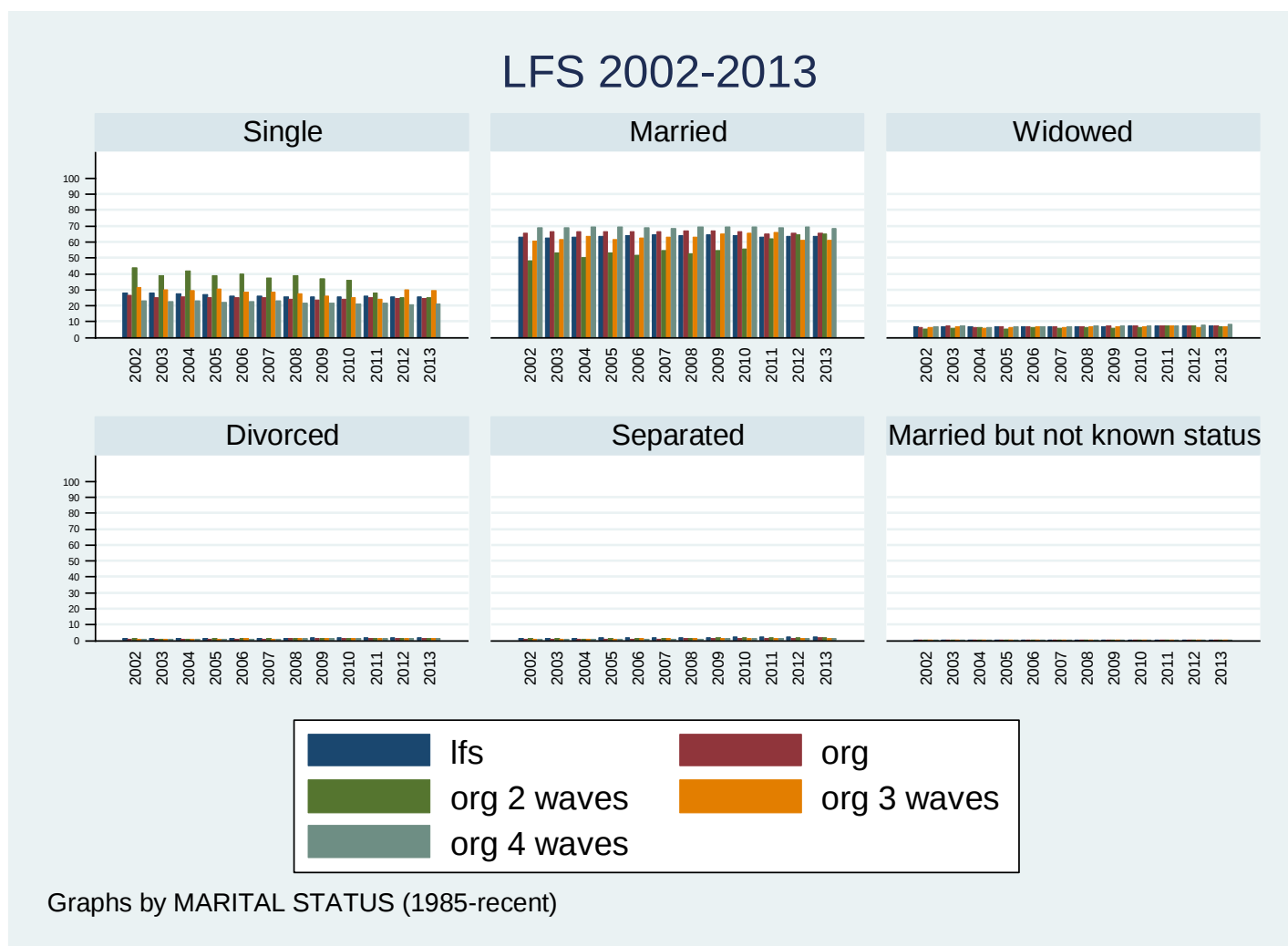
จากรูปที่ 54 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่ญาติอื่น ๆ พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มคงที่ ยกเว้นข้อมูล ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 2 ครั้ง ที่แนวโน้มลดลงอย่างชัดเจน โดยจะเห็นได้ว่าร้อยละเฉลี่ยของคนที่เป็นญาติอื่น ๆ จากข้อมูล ORG 2 waves มีค่าสูงที่สุดอยู่ที่ร้อยละ 4-13 ในขณะที่ข้อมูล LFS มีร้อยละเฉลี่ยค่อนข้างคงที่อยู่ที่ร้อยละ 4.8-5.0 และข้อมูล ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 3 ครั้ง มีค่าร้อยละเฉลี่ยอยู่ที่ 3-6 และข้อมูล ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 4 ครั้ง waves มีร้อยละเฉลี่ยอยู่ที่ 2.1-2.5 ตามลำดับ โดยที่ข้อมูล ORG ที่คนกลุ่มตัวอย่างซ้ำถูกสัมภาษณ์ 3 ครั้ง จะมีร้อยละเฉลี่ยของคนที่เป็นญาติอื่น ๆ ใกล้เคียงกับข้อมูล LFS มากที่สุด



รูปที่ 55: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่เป็นผู้อาศัย และคนรับใช้

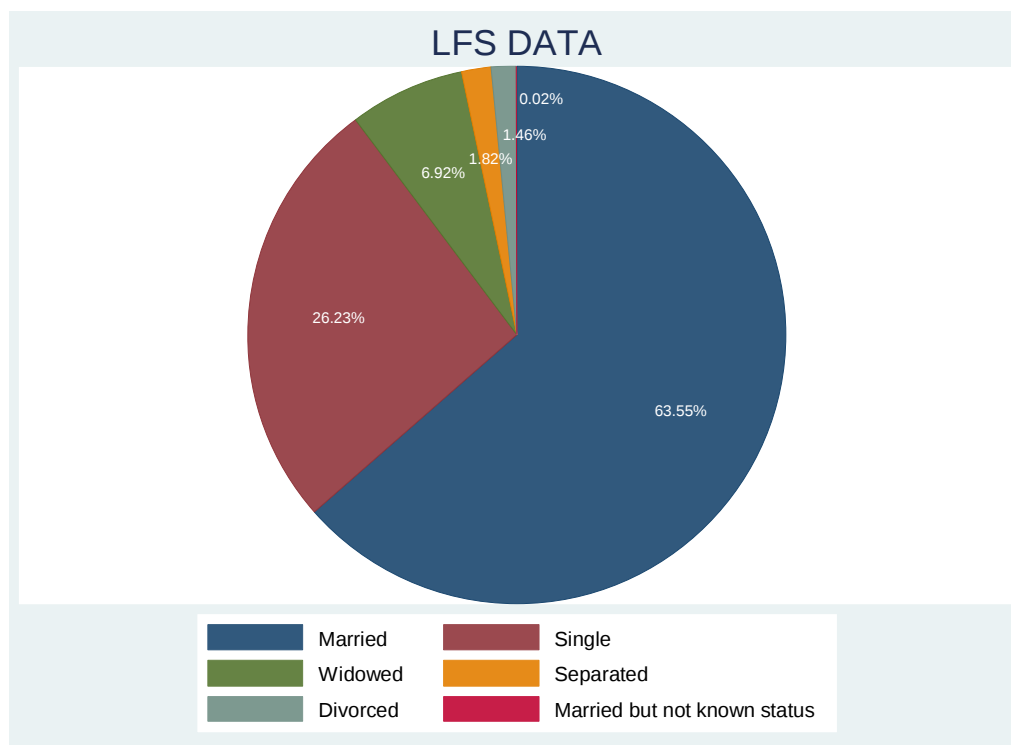
จากรูปที่ 55 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่เป็นผู้อาศัย และคนรับใช้ พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มลดลงอย่างต่อเนื่อง โดยมีคนที่เป็นผู้อาศัย และคนรับใช้เฉลี่ยร้อยละ 0.4-2 ซึ่งพบว่าข้อมูลตัวอย่างซ้ำที่มีร้อยละโดยเฉลี่ยของผู้อาศัย และคนรับใช้ใกล้เคียงกับข้อมูล LFS มากที่สุดคือข้อมูล ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 3 ครั้ง

เมื่อพิจารณาร้อยละของจำนวนกลุ่มตัวอย่างที่แบ่งตามสถานภาพ จากรูปที่ 55 พบว่า ทั้งข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีร้อยละของคนที่ตั้งงานแล้วมากที่สุด คิดเป็นร้อยละ 61-69 รองลงมาเป็นคนที่ยังโสด คิดเป็นร้อยละ 24-48 และลำดับที่ 3 คือคนที่เป็นม่าย คิดเป็นร้อยละ 7 และลำดับที่ 4 คือคนที่หย่าร้าง คิดเป็นร้อยละ 1-2 และลำดับที่ 5 คือคนที่แยกกันอยู่ คิดเป็นร้อยละ 1-2 และลำดับสุดท้ายคือ คนที่เคยสมรสแต่ไม่ทราบสถานภาพ คิดเป็นร้อยละ 0.01 ตามลำดับ



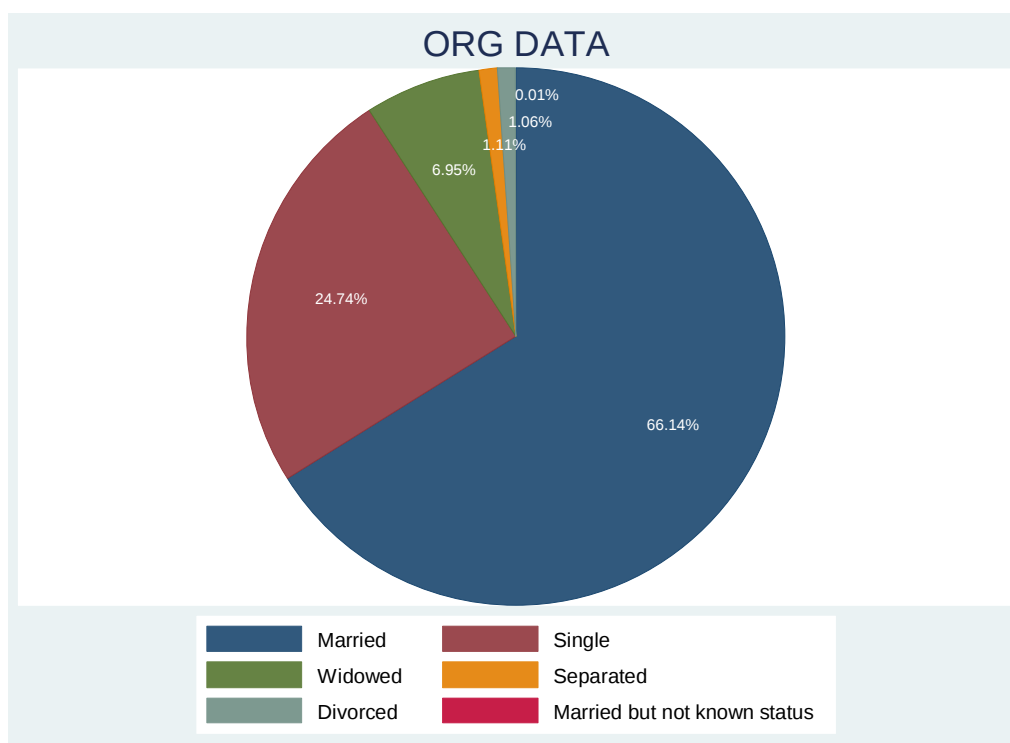
รูปที่ 56: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG แบ่งตามสถานภาพ

โดยที่ร้อยละของคนตามสถานภาพในกลุ่มต่าง ๆ แยกตามชุดข้อมูล ได้แก่ LFS, ORG, ORG กลุ่มตัวอย่างถูกสัมภาษณ์ 2 ครั้ง, 3 ครั้ง และ 4 ครั้ง ตามลำดับ สามารถแสดงได้ในรูปที่ 57 - 61 ดังต่อไปนี้



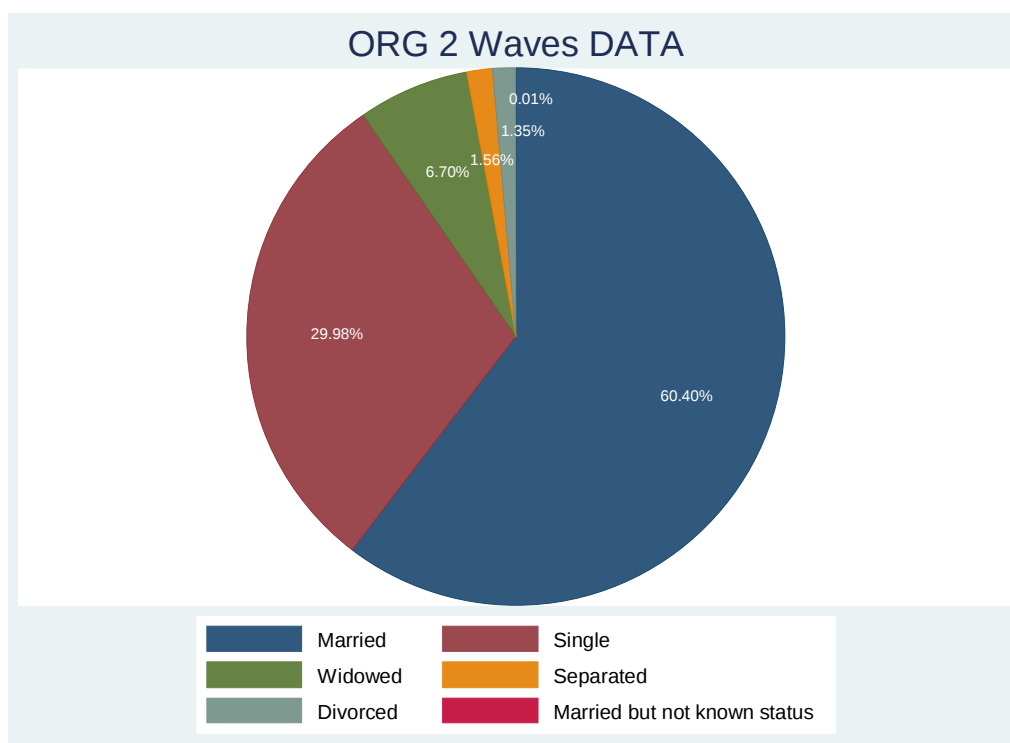
รูปที่ 57: ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS ตั้งแต่ปี 2002-2013
แบ่งตามสถานภาพ

จากรูปที่ 57 แสดงร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างที่แบ่งตามสถานภาพจากข้อมูล LFS ตั้งแต่ปี 2002-2013 พบว่า มีร้อยละของคนที่แต่งงานแล้วมากที่สุด คิดเป็นร้อยละ 63.55 รองลงมาเป็นคนที่ยังโสด คิดเป็นร้อยละ 26.23 และลำดับที่ 3 คือคนที่เป็นม่าย คิดเป็นร้อยละ 6.92 และลำดับที่ 4 คือคนที่หย่าร้าง คิดเป็นร้อยละ 1.82 และลำดับที่ 5 คือคนที่แยกกันอยู่ คิดเป็นร้อยละ 1.46 และลำดับสุดท้ายคือ คนที่เคยสมรสแต่ไม่ทราบสถานภาพ คิดเป็นร้อยละ 0.02 ตามลำดับ



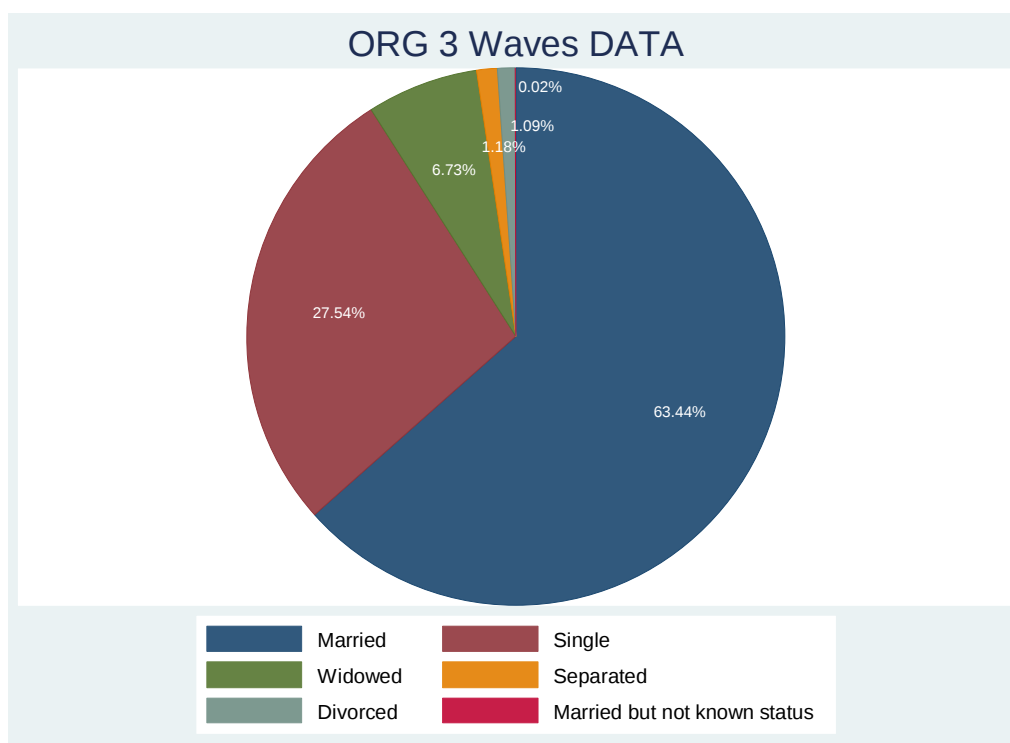
รูปที่ 58: ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG ตั้งแต่ปี 2002-2013
แบ่งตามสถานภาพ

จากรูปที่ 58 แสดงร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างที่แบ่งตามสถานภาพจากข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์มากกว่า 1 ครั้งขึ้นไป ตั้งแต่ปี 2002-2013 พบว่า มีร้อยละของคนที่แต่งงานแล้วมากที่สุด คิดเป็นร้อยละ 63.67 รองลงมาเป็นคนที่ยังโสด คิดเป็นร้อยละ 24.46 และลำดับที่ 3 คือคนที่เป้นม่าย คิดเป็นร้อยละ 6.91 และลำดับที่ 4 คือคนที่หย่าร้าง คิดเป็นร้อยละ 1.60 และลำดับที่ 5 คือคนที่แยกกันอยู่ คิดเป็นร้อยละ 1.35 และลำดับสุดท้ายคือ คนที่เคยสมรสแต่ไม่ทราบสถานภาพ คิดเป็นร้อยละ 0.01 ตามลำดับ



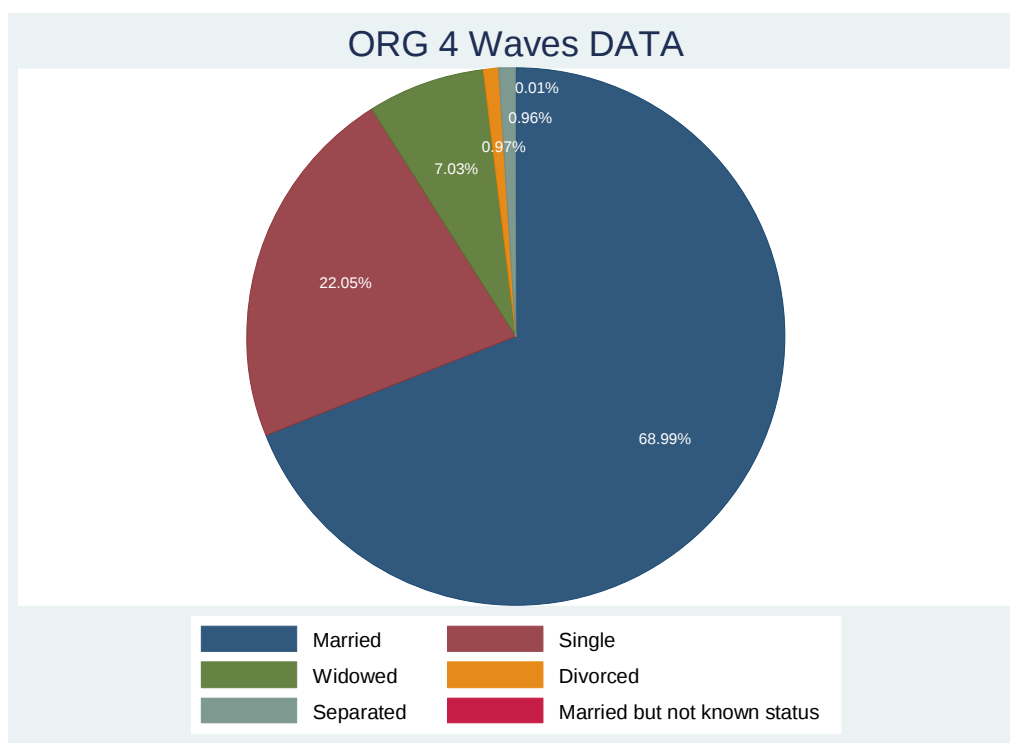
รูปที่ 59: ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 2 waves ตั้งแต่ปี 2002-2013 แบ่งตามสถานภาพ

จากรูปที่ 59 แสดงร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างที่แบ่งตามสถานภาพจากข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์ 2 ครั้ง ตั้งแต่ปี 2002-2013 พบว่า มีร้อยละของคนที่แต่งงานแล้วมากที่สุด คิดเป็นร้อยละ 60.40 รองลงมาเป็นคนที่ยังโสด คิดเป็นร้อยละ 29.98 และลำดับที่ 3 คือคนที่เป็นม่าย คิดเป็นร้อยละ 6.70 และลำดับที่ 4 คือคนที่หย่าร้าง คิดเป็นร้อยละ 1.56 และลำดับที่ 5 คือคนที่แยกกันอยู่ คิดเป็นร้อยละ 1.35 และลำดับสุดท้ายคือ คนที่เคยสมรสแต่ไม่ทราบสถานภาพ คิดเป็นร้อยละ 0.01 ตามลำดับ



รูปที่ 60: ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 3 waves ตั้งแต่ปี 2002-2013 แบ่งตามสถานภาพ

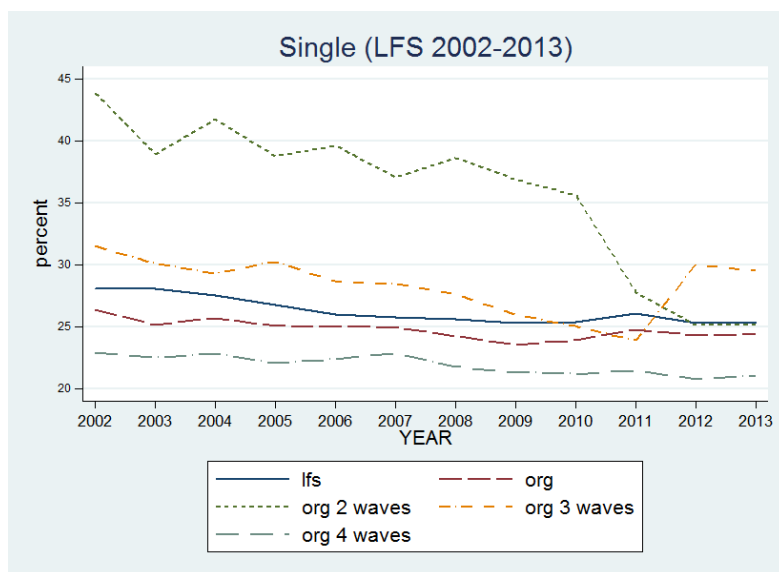
จากรูปที่ 60 แสดงร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างที่แบ่งตามสถานภาพจากข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์ 3 ครั้ง ตั้งแต่ปี 2002-2013 พบว่า มีร้อยละของคนที่แต่งงานแล้วมากที่สุด คิดเป็นร้อยละ 63.44 รองลงมาเป็นคนที่ยังโสด คิดเป็นร้อยละ 27.54 และลำดับที่ 3 คือคนที่เป้นม่าย คิดเป็นร้อยละ 6.73 และลำดับที่ 4 คือคนที่หย่าร้าง คิดเป็นร้อยละ 1.18 และลำดับที่ 5 คือคนที่แยกกันอยู่ คิดเป็นร้อยละ 1.09 และลำดับสุดท้ายคือ คนที่เคยสมรสแต่ไม่ทราบสถานภาพ คิดเป็นร้อยละ 0.02 ตามลำดับ



รูปที่ 61: ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 4 waves ตั้งแต่ปี 2002-2013 แบ่งตามสถานภาพ

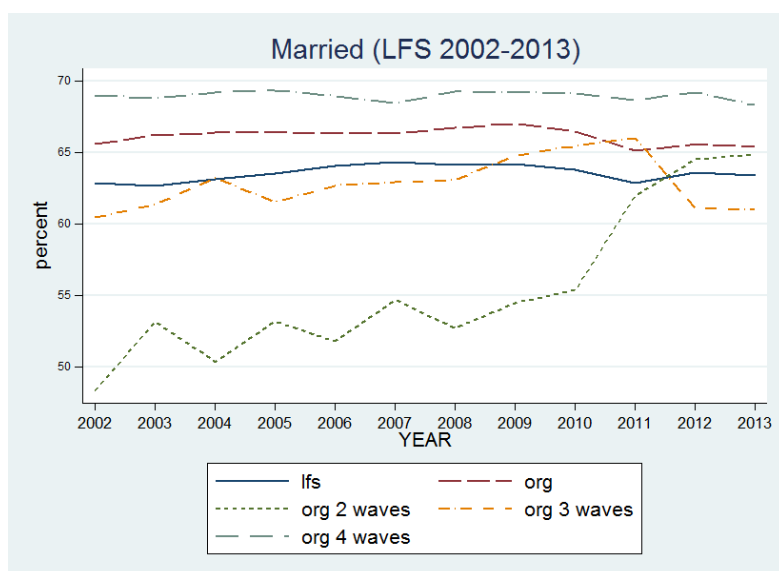
จากรูปที่ 61 แสดงร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างที่แบ่งตามสถานภาพจากข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์ 4 ครั้ง ตั้งแต่ปี 2002-2013 พบว่า มีร้อยละของคนที่แต่งงานแล้วมากที่สุด คิดเป็นร้อยละ 68.99 รองลงมาเป็นคนที่ยังโสด คิดเป็นร้อยละ 22.05 และลำดับที่ 3 คือคนที่เป็นม่าย คิดเป็นร้อยละ 7.03 และลำดับที่ 4 คือคนที่หย่าร้าง คิดเป็นร้อยละ 0.97 และลำดับที่ 5 คือคนที่แยกกันอยู่ คิดเป็นร้อยละ 0.96 และลำดับสุดท้ายคือ คนที่เคยสมรสแต่ไม่ทราบสถานภาพ คิดเป็นร้อยละ 0.01 ตามลำดับ

ส่วนถัดไปจะเป็นการเปรียบเทียบกราฟเส้นที่แสดงร้อยละโดยเฉลี่ยต่อปีของแต่ละชุดข้อมูลที่แบ่งตามสถานภาพ ซึ่งสามารถแสดงได้ดังต่อไปนี้



รูปที่ 62: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG สถานภาพโสด

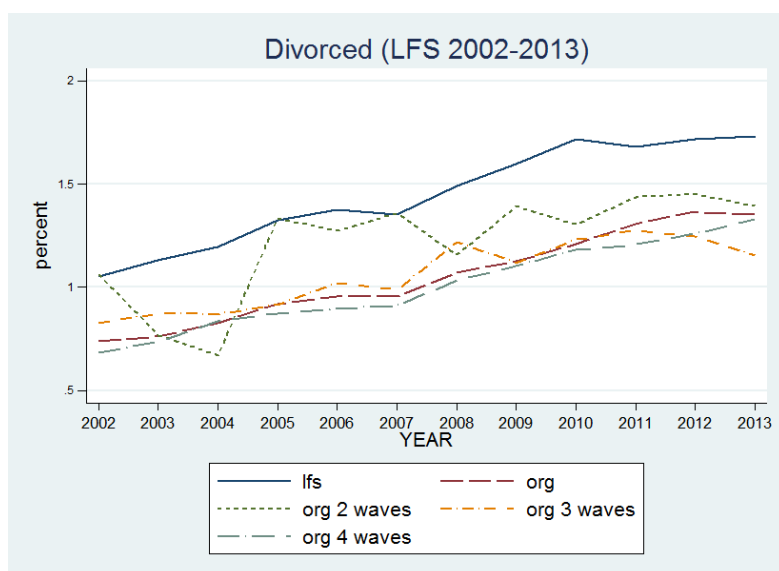
จากรูปที่ 62 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่มีสถานภาพโสด พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มคงที่ โดยมีคนที่โสดเฉลี่ยร้อยละ 22-28 ยกเว้นข้อมูลตัวอย่างซ้ำ ORG 2 waves ที่มีร้อยละเฉลี่ยของคนโสดสูงกว่าข้อมูลชุดอื่นๆ ซึ่งมีค่าเฉลี่ยร้อยละ 24-44 ซึ่งจากการเปรียบเทียบข้อมูลตัวอย่างซ้ำ ORG กับข้อมูล LFS พบว่าข้อมูลตัวอย่างซ้ำที่มีร้อยละโดยเฉลี่ยของคนโสดใกล้เคียงกับข้อมูล LFS มากที่สุดคือ ข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 3 ครั้ง



รูปที่ 63: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG สถานภาพสมรส

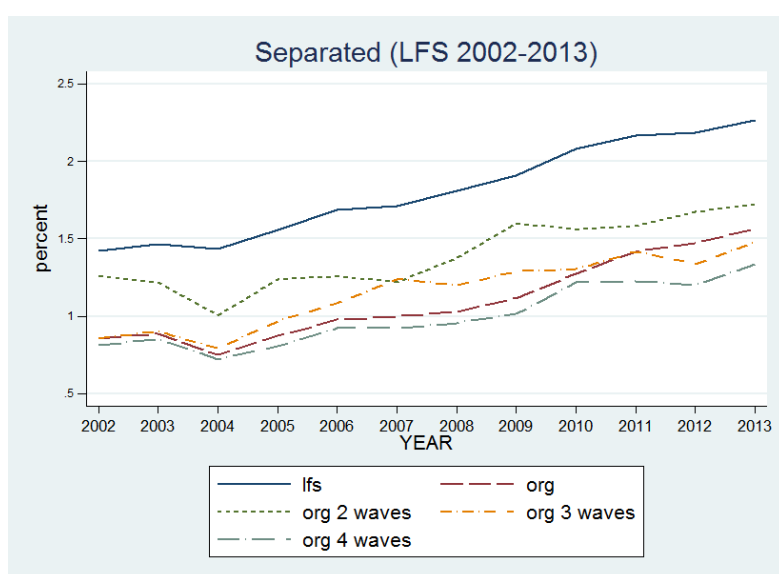
จากรูปที่ 63 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่มีสถานภาพสมรส พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มคงที่ โดยมีคนที่สมรสเฉลี่ยร้อยละ 60-70 ยกเว้นข้อมูลตัวอย่างซ้ำ ORG 2

waves ที่มีร้อยละเฉลี่ยของคนที่มีสมรสต่ำกว่าข้อมูลชุดอื่นๆ ซึ่งมีค่าเฉลี่ยร้อยละ 45-65 ซึ่งจากการเปรียบเทียบข้อมูลตัวอย่างซ้ำ ORG กับข้อมูล LFS พบว่าข้อมูลตัวอย่างซ้ำที่มีร้อยละโดยเฉลี่ยของคนที่มีสมรสใกล้เคียงกับข้อมูล LFS มากที่สุดคือ ข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง



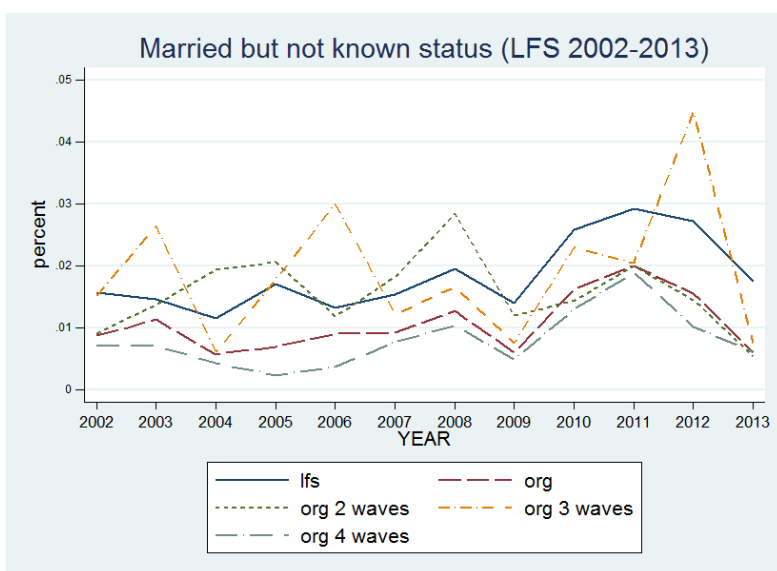
รูปที่ 64: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG สถานภาพหย่า

จากรูปที่ 64 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่มีสถานภาพหย่า พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มเพิ่มขึ้นอย่างต่อเนื่อง โดยคนที่มีสถานภาพหย่ามีค่าเฉลี่ยร้อยละ 0.8-1.8 ซึ่งจากการเปรียบเทียบข้อมูลตัวอย่างซ้ำ ORG กับข้อมูล LFS พบว่าข้อมูลตัวอย่างซ้ำที่มีร้อยละโดยเฉลี่ยของคนที่มีสถานภาพหย่าใกล้เคียงกับข้อมูล LFS มากที่สุดคือ ข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 2 ครั้ง



รูปที่ 65: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG สถานภาพแยกกันอยู่

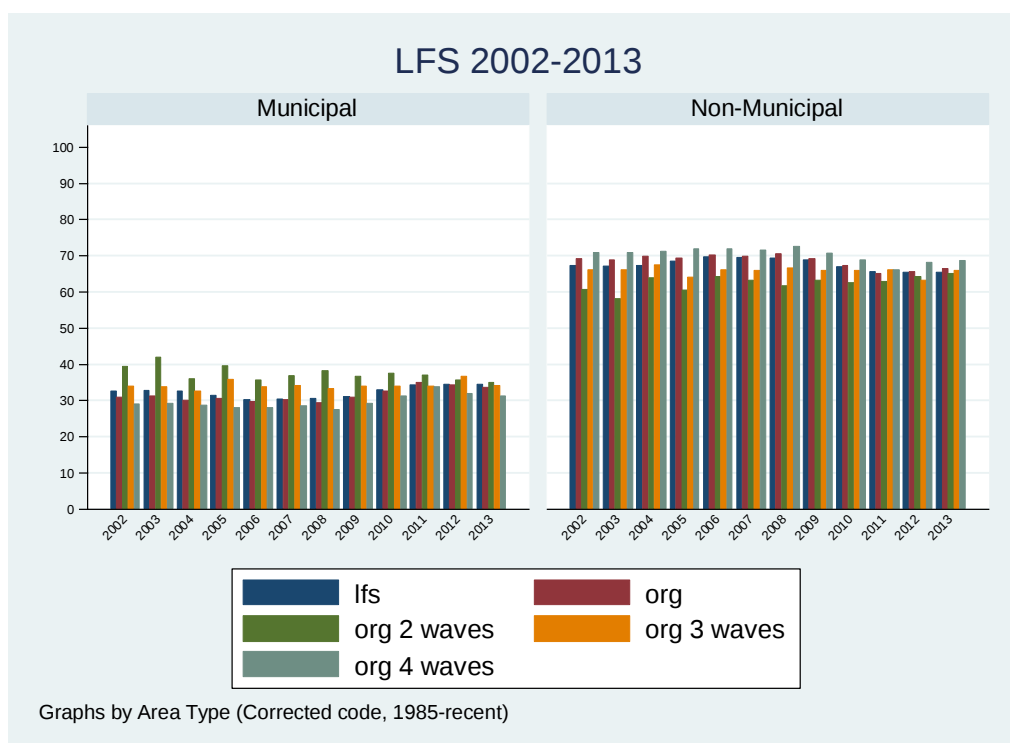
จากรูปที่ 65 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่มีสถานภาพแยกกันอยู่ พบว่า ข้อมูล LFS และ ข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มเพิ่มขึ้นอย่างต่อเนื่อง โดยคนที่มีสถานภาพแยกกันอยู่มีค่าเฉลี่ยร้อยละ 0.7-2.3 ซึ่งจากการเปรียบเทียบข้อมูลตัวอย่างซ้ำ ORG กับข้อมูล LFS พบว่าข้อมูลตัวอย่างซ้ำที่มีร้อยละโดยเฉลี่ยของคนที่มีสถานภาพแยกกันใกล้เคียงกับข้อมูล LFS มากที่สุดคือ ข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 2 ครั้ง



รูปที่ 66: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่เคยสมรสแต่ไม่ทราบสถานภาพ

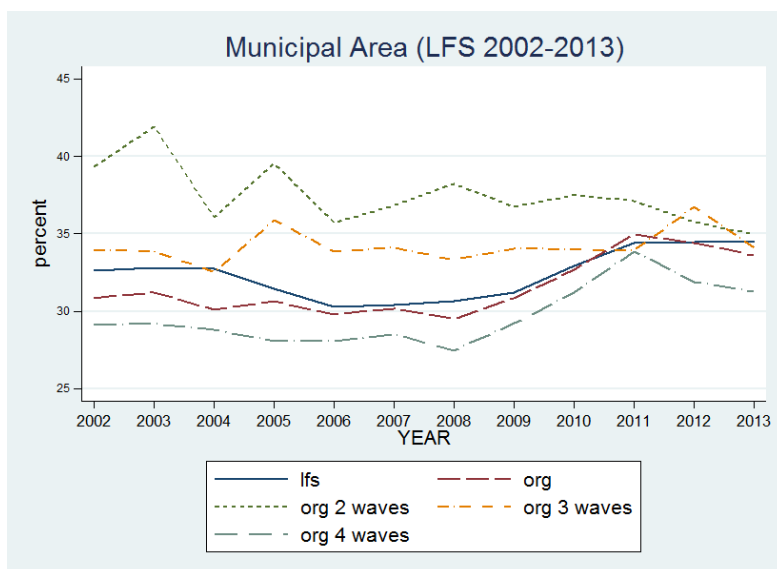
จากรูปที่ 66 แสดงร้อยละของจำนวนกลุ่มตัวอย่างที่เคยสมรสแต่ไม่ทราบสถานภาพ พบว่า ข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG ไม่มีรูปแบบที่แน่นอนมีทั้งเพิ่มขึ้นและลดลง โดยคนที่มีเคยสมรสแต่ไม่ทราบสถานภาพมีค่าเฉลี่ยร้อยละ 0.01-0.05 ซึ่งเป็นร้อยละที่ค่อนข้างน้อยมาก โดยเมื่อเปรียบเทียบข้อมูลตัวอย่างซ้ำ ORG ที่มีร้อยละของคนที่เคยสมรสแต่ไม่ทราบสถานภาพซึ่งมีค่าใกล้เคียงกับข้อมูล LFS มากที่สุดคือข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 2 ครั้ง

เมื่อพิจารณาร้อยละของจำนวนกลุ่มตัวอย่างรายบุคคลที่แบ่งตามเขตที่อยู่อาศัย จากรูปที่ 67 พบว่า ทั้งข้อมูล LFS และ ORG มีร้อยละของคนที่มีอายุอยู่นอกเขตเทศบาลสูงกว่าคนที่อาศัยอยู่ในเขตเทศบาล โดยเฉลี่ยคิดเป็นคนที่อาศัยอยู่ในเขตเทศบาลร้อยละ 30 และคนที่อาศัยอยู่นอกเขตเทศบาลร้อยละ 70



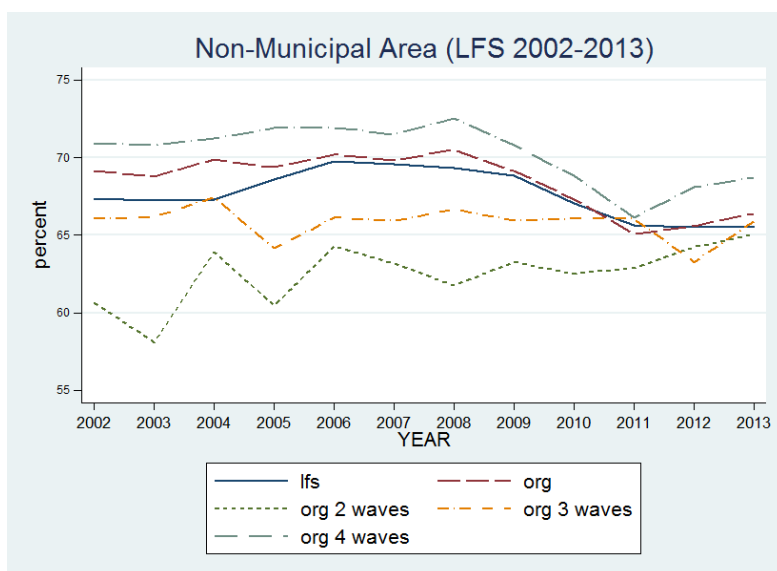
รูปที่ 67: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG แบ่งตามเขตที่อยู่อาศัย

และเมื่อเปรียบเทียบกราฟเส้นที่แสดงร้อยละโดยเฉลี่ยของคนที่อยู่ในเขตเทศบาลและนอกเขตเทศบาลจากข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG ซึ่งสามารถแสดงได้จากรูปที่ 68-69 ดังนี้



รูปที่ 68: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่อยู่อาศัยในเขตเทศบาล

จากรูปที่ 68 พบว่า ทั้งจากข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มที่ร้อยละของคน ที่อาศัยอยู่ในเขตเทศบาลลดลงในช่วงปี 2002-2008 แต่กลับมีแนวโน้มเพิ่มขึ้นในช่วงปี 2009-2013 ซึ่งโดยเฉลี่ยมีคน ที่อาศัยอยู่ในเขตเทศบาลร้อยละ 27-40 โดยร้อยละของคน ที่อยู่ในเขตเทศบาลจากข้อมูล ORG กลุ่มตัวอย่าง ซ้ำที่ถูกสัมภาษณ์ 3 ครั้ง มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด

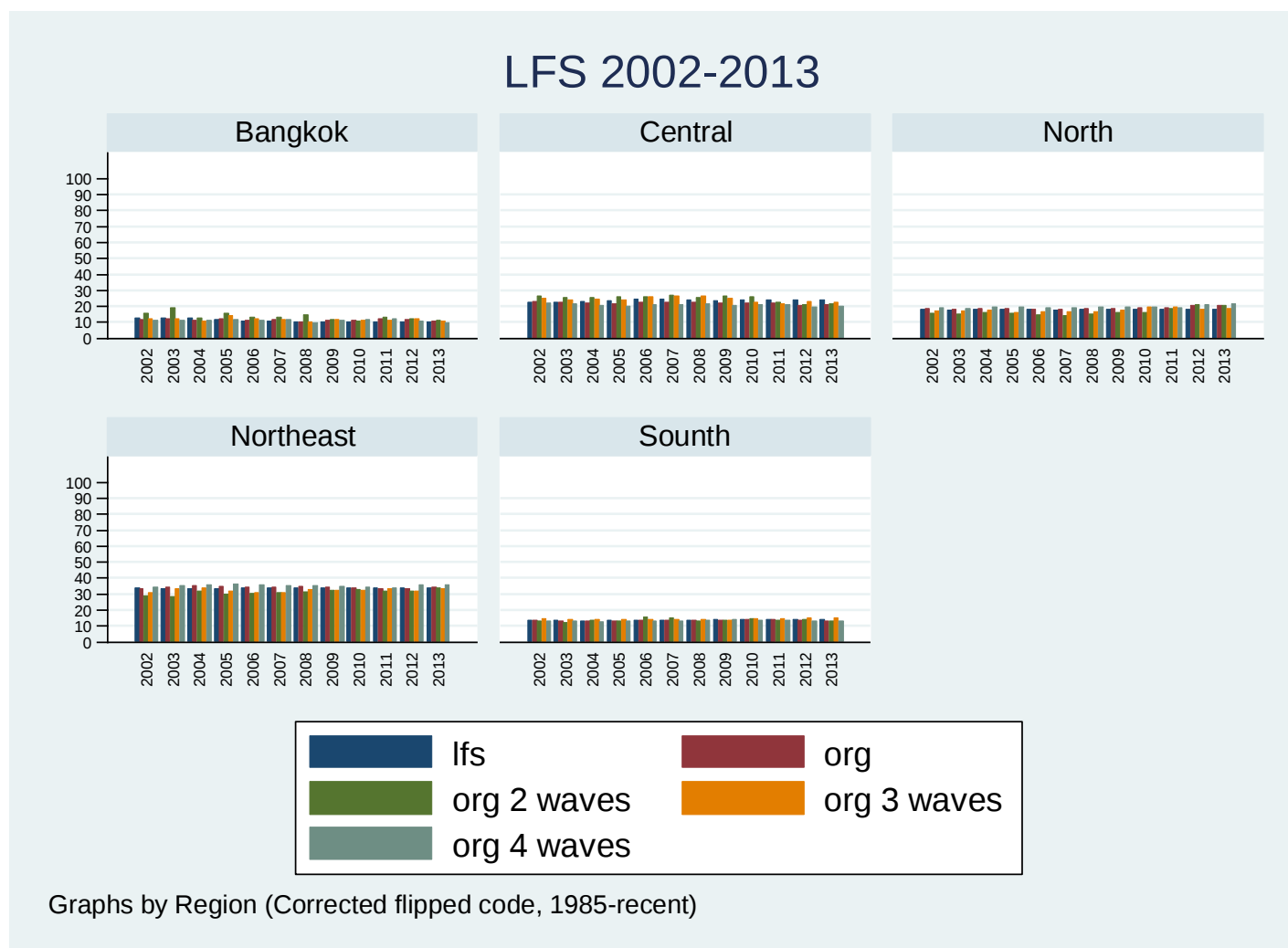


รูปที่ 69: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่อยู่นอกเขตเทศบาล

จากรูปที่ 69 พบว่า ทั้งจากข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มที่ร้อยละของคน ที่อาศัยอยู่นอกเขตเทศบาลเพิ่มขึ้นในช่วงปี 2002-2006 แต่กลับมีแนวโน้มลดลงในช่วงปี 2007-2013 ซึ่งคนที่อาศัย

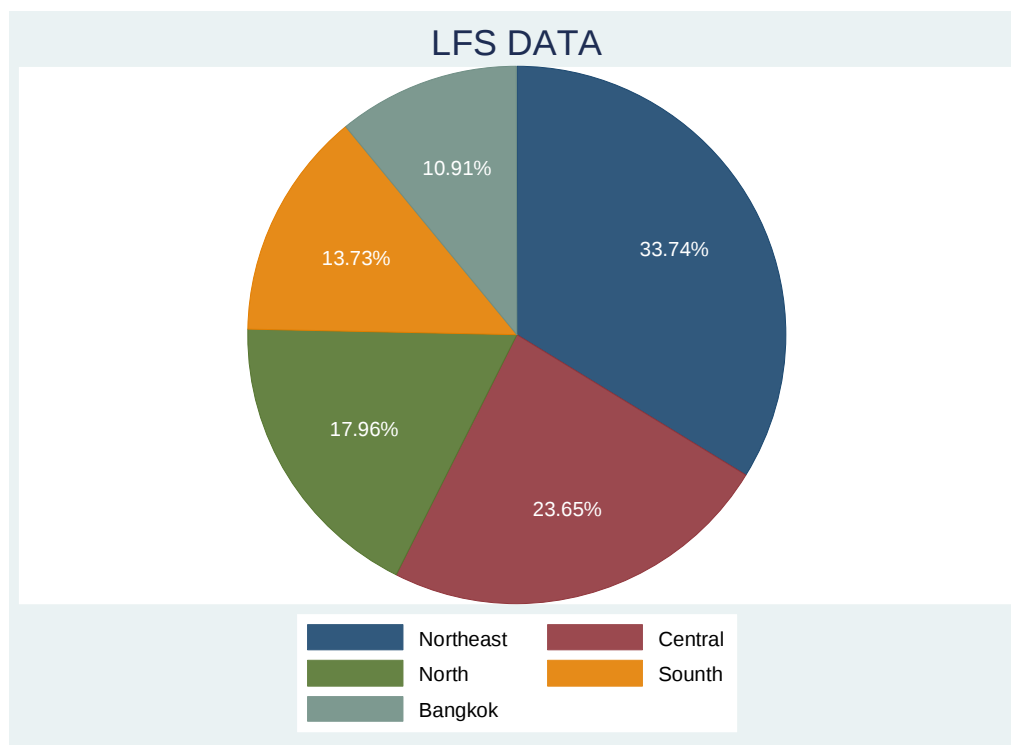
อยู่นอกเขตเทศบาลมีจำนวนเฉลี่ยร้อยละ 59-73 โดยร้อยละของคนที่อยู่นอกเขตเทศบาลจากข้อมูล ORG ที่กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์มากกว่า 1 ครั้งมีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด

เมื่อพิจารณาร้อยละของจำนวนกลุ่มตัวอย่างที่แบ่งตามภาค จากรูปที่ 70 พบว่า ทั้งข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีร้อยละของคนที่อยู่ในภาคตะวันออกเฉียงเหนือมากที่สุด คิดเป็นร้อยละ 40 รองลงมาเป็นคนที่อยู่ในภาคกลาง คิดเป็นร้อยละ 30 และลำดับที่ 3 คือคนที่อยู่ในภาคเหนือ คิดเป็นร้อยละ 20 และลำดับที่ 4 คือคนที่อยู่ในภาคใต้ คิดเป็นร้อยละ 18 และลำดับสุดท้ายคือคนที่อยู่กรุงเทพฯ คิดเป็นร้อยละ 12 ตามลำดับ



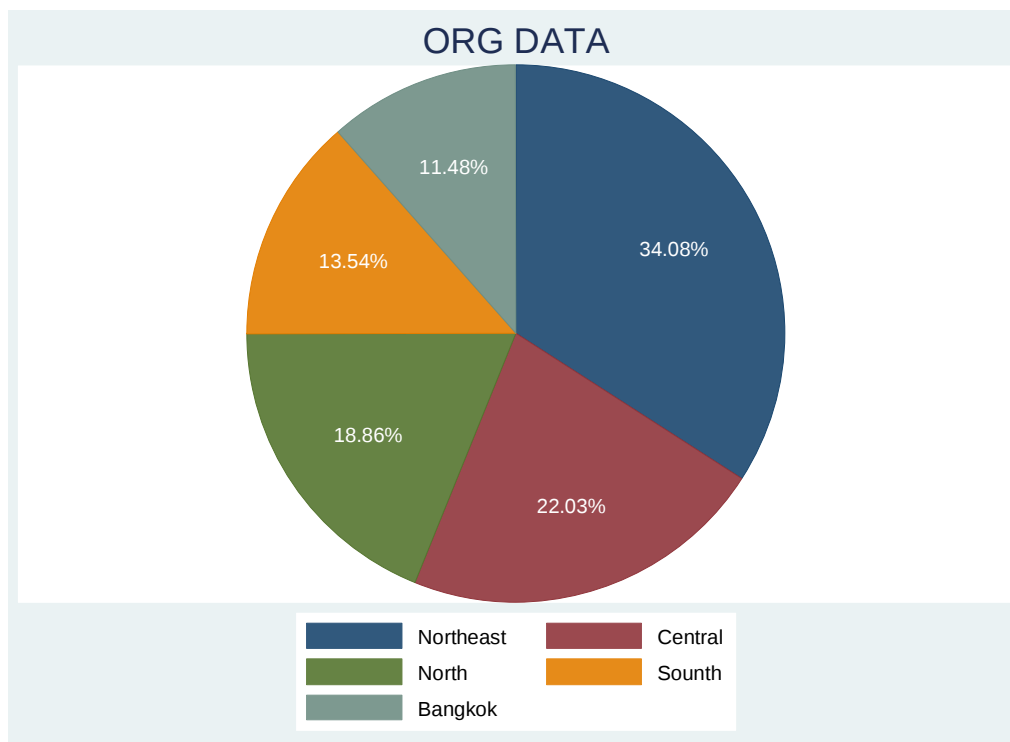
รูปที่ 70: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG แบ่งตามภาค

โดยร้อยละของคนกลุ่มตัวอย่างที่แบ่งตามภาค แยกตามชุดข้อมูล ได้แก่ LFS, ORG, ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 2 ครั้ง, 3 ครั้ง และ 4 ครั้ง ตามลำดับ สามารถแสดงได้ในรูปที่ 71-75 ดังต่อไปนี้



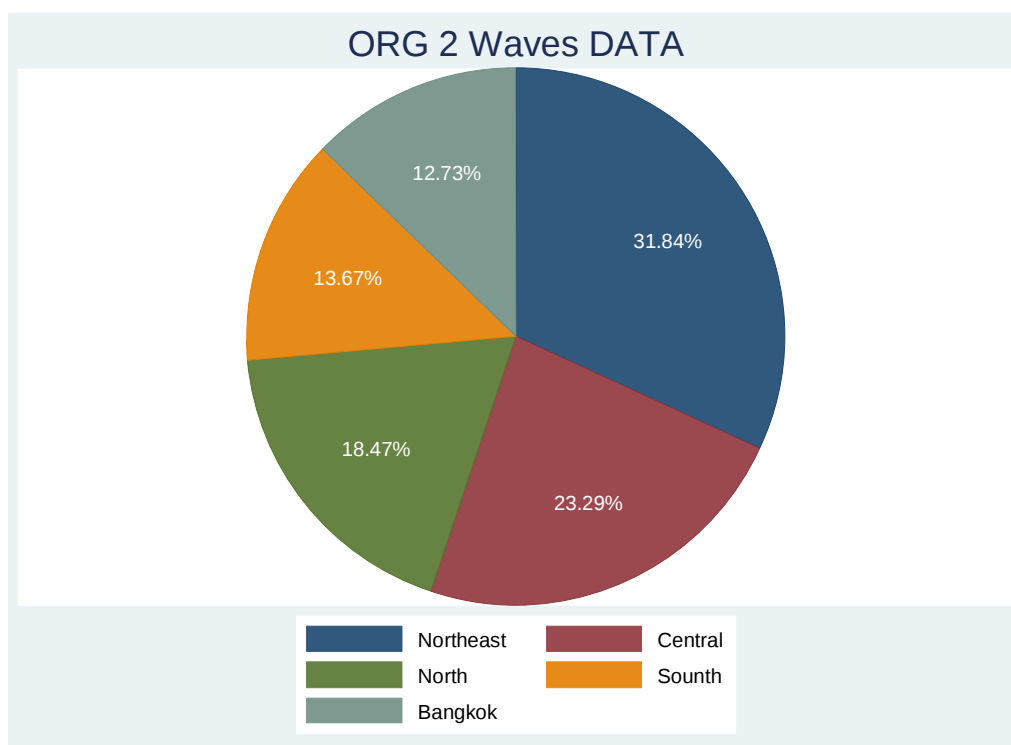
รูปที่ 71: ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS ตั้งแต่ปี 2002-2013 แบ่งตามภาค

จากรูปที่ 71 แสดงร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างที่แบ่งตามภาคจากข้อมูล LFS ตั้งแต่ปี 2002-2013 พบว่า มีร้อยละของคนที่อยู่ในภาคตะวันออกเฉียงเหนือมากที่สุด เฉลี่ยร้อยละ 33.74 รองลงมาเป็นคนที่อยู่ในภาคกลาง เฉลี่ยร้อยละ 23.65 และลำดับที่ 3 คือคนที่อยู่ในภาคเหนือ เฉลี่ยร้อยละ 17.96 และลำดับที่ 4 คือคนที่อยู่ในภาคใต้ เฉลี่ยร้อยละ 13.73 และลำดับสุดท้ายคือคนที่อยู่กรุงเทพฯ เฉลี่ยร้อยละ 10.91 ตามลำดับ



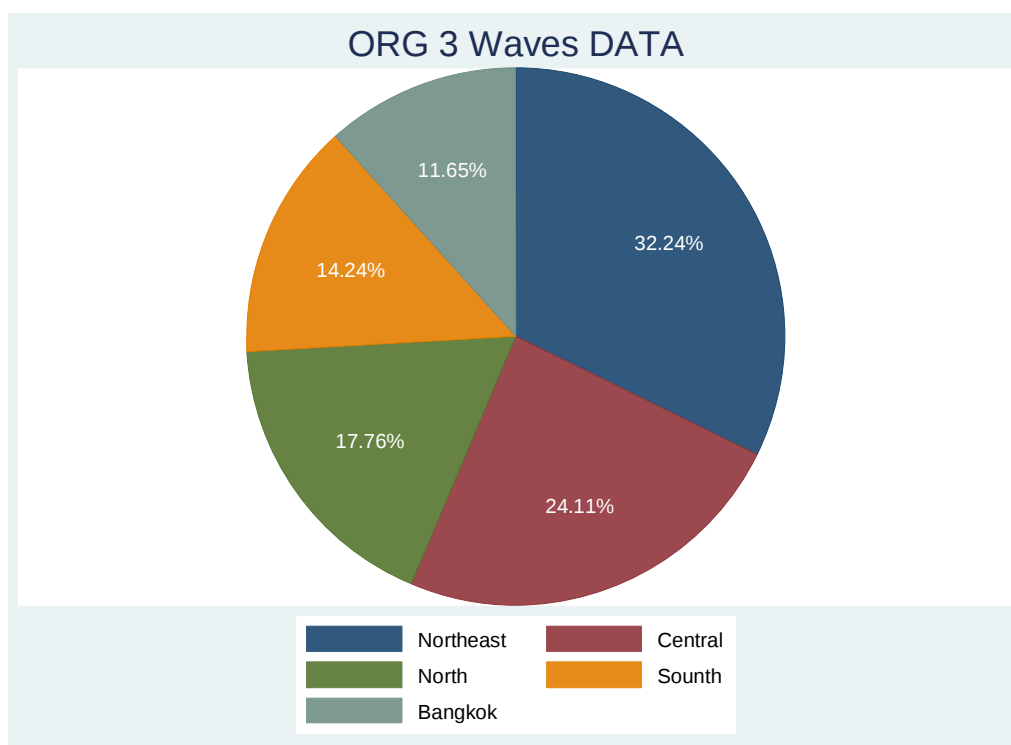
รูปที่ 72: ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG ตั้งแต่ปี 2002-2013 แบ่งตามภาค

จากรูปที่ 72 แสดงร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างที่แบ่งตามภาคจากข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์มากกว่า 1 ครั้งขึ้นไป ตั้งแต่ปี 2002-2013 พบว่า มีร้อยละของคนที่อยู่ในภาคตะวันออกเฉียงเหนือมากที่สุด เฉลี่ยร้อยละ 34.08 รองลงมาเป็นคนที่อยู่ในภาคกลาง เฉลี่ยร้อยละ 22.03 และลำดับที่ 3 คือคนที่อยู่ในภาคเหนือ เฉลี่ยร้อยละ 18.86 และลำดับที่ 4 คือคนที่อยู่ในภาคใต้ เฉลี่ยร้อยละ 13.54 และลำดับสุดท้ายคือคนที่อยู่กรุงเทพฯ เฉลี่ยร้อยละ 11.48 ตามลำดับ



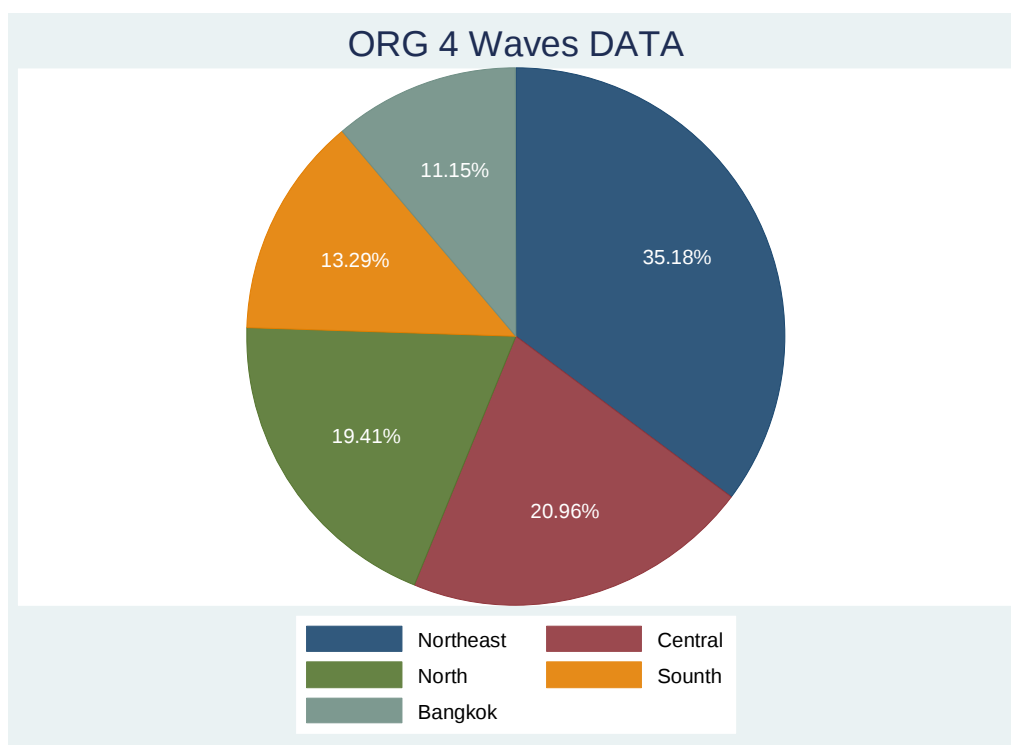
รูปที่ 73: ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 2 waves ตั้งแต่ปี 2002-2013
แบ่งตามภาค

จากรูปที่ 73 แสดงร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างที่แบ่งตามภาคจากข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์ 2 ครั้ง ตั้งแต่ปี 2002-2013 พบว่า มีร้อยละของคนที่อยู่ในภาคตะวันออกเฉียงเหนือมากที่สุด เฉลี่ยร้อยละ 31.84 รองลงมาเป็นคนที่อยู่ในภาคกลาง เฉลี่ยร้อยละ 23.29 และลำดับที่ 3 คือคนที่อยู่ในภาคเหนือ เฉลี่ยร้อยละ 18.47 และลำดับที่ 4 คือคนที่อยู่ในภาคใต้ เฉลี่ยร้อยละ 13.67 และลำดับสุดท้ายคือคนที่อยู่กรุงเทพฯ เฉลี่ยร้อยละ 12.73 ตามลำดับ



รูปที่ 74: ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 3 waves ตั้งแต่ปี 2002-2013
แบ่งตามภาค

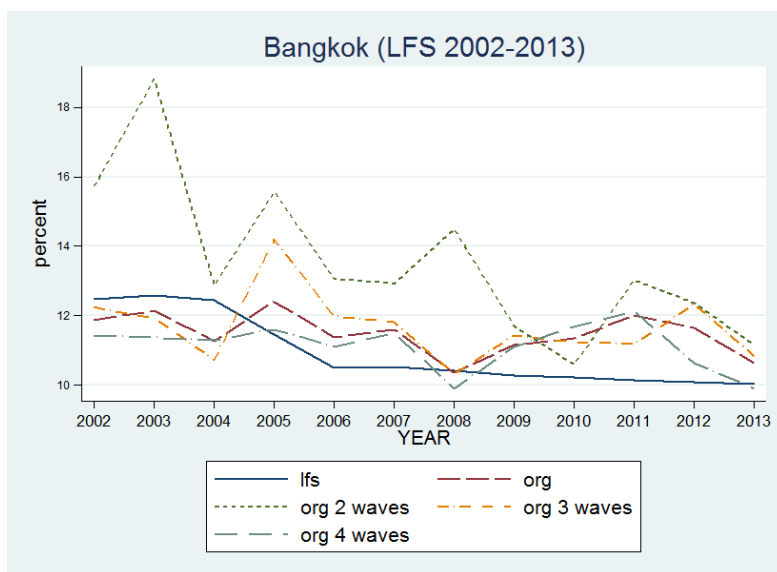
จากรูปที่ 74 แสดงร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างที่แบ่งตามภาคจากข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์ 3 ครั้ง ตั้งแต่ปี 2002-2013 พบว่า มีร้อยละของคนที่อยู่ในภาคตะวันออกเฉียงเหนือมากที่สุด เฉลี่ยร้อยละ 32.24 รองลงมาเป็นคนที่อยู่ในภาคกลาง เฉลี่ยร้อยละ 24.11 และลำดับที่ 3 คือคนที่อยู่ในภาคเหนือ เฉลี่ยร้อยละ 17.76 และลำดับที่ 4 คือคนที่อยู่ในภาคใต้ เฉลี่ยร้อยละ 14.24 และลำดับสุดท้ายคือคนที่อยู่กรุงเทพฯ เฉลี่ยร้อยละ 11.65 ตามลำดับ



รูปที่ 75: ร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างจากข้อมูล ORG 4 waves ตั้งแต่ปี 2002-2013
แบ่งตามภาค

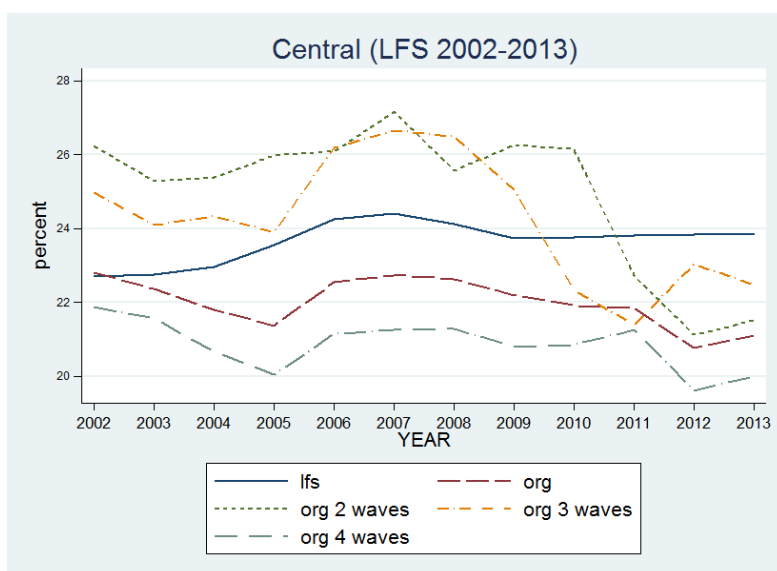
จากรูปที่ 75 แสดงร้อยละโดยเฉลี่ยของจำนวนกลุ่มตัวอย่างที่แบ่งตามภาคจากข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์ 4 ครั้ง ตั้งแต่ปี 2002-2013 พบว่า มีร้อยละของคนที่อยู่ในภาคตะวันออกเฉียงเหนือมากที่สุด เฉลี่ยร้อยละ 35.18 รองลงมาเป็นคนที่อยู่ในภาคกลาง เฉลี่ยร้อยละ 20.96 และลำดับที่ 3 คือคนที่อยู่ในภาคเหนือ เฉลี่ยร้อยละ 19.41 และลำดับที่ 4 คือคนที่อยู่ในภาคใต้ เฉลี่ยร้อยละ 13.29 และลำดับสุดท้ายคือคนที่อยู่กรุงเทพฯ เฉลี่ยร้อยละ 11.15 ตามลำดับ

ส่วนถัดไปจะเป็นการเปรียบเทียบกราฟเส้นที่แสดงร้อยละโดยเฉลี่ยต่อปีของแต่ละชุดข้อมูลที่แบ่งตามภาค ซึ่งสามารถแสดงได้ดังต่อไปนี้



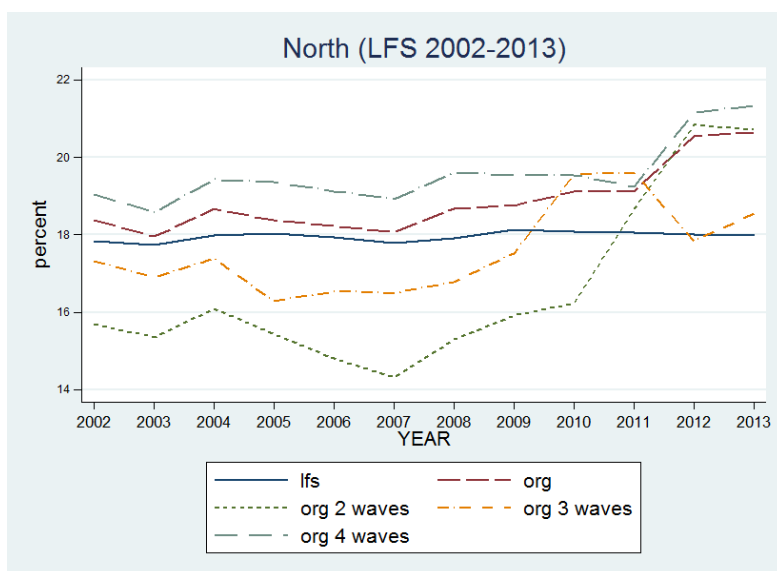
รูปที่ 76: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่อยู่กรุงเทพฯ

จากรูปที่ 76 แสดงร้อยละกลุ่มตัวอย่างที่อยู่ในกรุงเทพฯ พบว่า ข้อมูล LFS มีร้อยละของคนที่อยู่ในกรุงเทพฯ ลดลงในช่วงปี 2002-2006 และหลังจากนั้นตั้งแต่ปี 2007-2013 มีร้อยละของคนที่อยู่ในกรุงเทพฯ ค่อนข้างคงที่ เมื่อเทียบกับข้อมูลตัวอย่างซ้ำ ORG โดยจะเห็นได้ว่า ข้อมูล ORG แต่ละชุดไม่มีรูปแบบที่แน่นอน และมีการกระจายตัวสูงกว่าข้อมูล LFS ซึ่งหากเปรียบเทียบข้อมูลตัวอย่างซ้ำ ORG ที่ถูกสัมภาษณ์ 4 ครั้ง น่าจะมีร้อยละของคนที่อยู่ในกรุงเทพฯ ใกล้เคียงกับข้อมูล LFS มากที่สุด



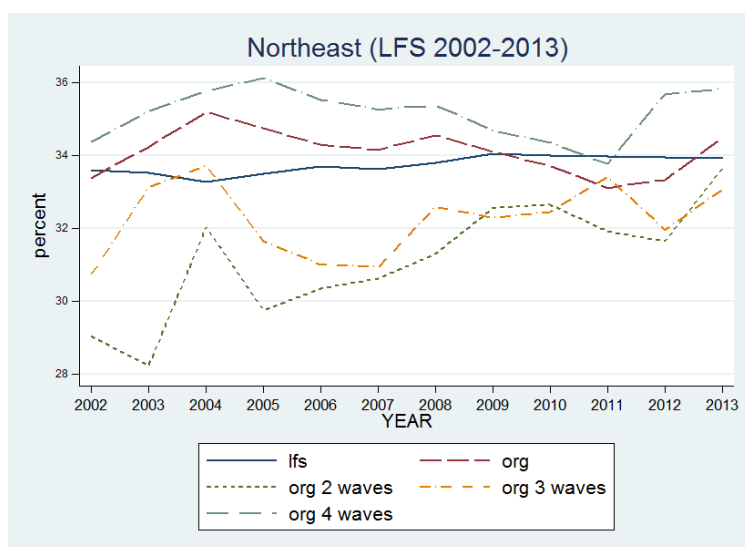
รูปที่ 77: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่อยู่ภาคกลาง

จากรูปที่ 77 แสดงร้อยละกลุ่มตัวอย่างที่อยู่ภาคกลาง พบว่า ข้อมูล LFS ร้อยละของคนที่อยู่ในภาคกลาง มีแนวโน้มเพิ่มขึ้นอย่างช้า ๆ ในขณะที่ข้อมูลตัวอย่างซ้ำ ORG กลับมีแนวโน้มที่ลดลง ซึ่งหากจะเปรียบเทียบร้อยละของคนที่อยู่ในภาคกลางที่มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด คือข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 3 ครั้ง



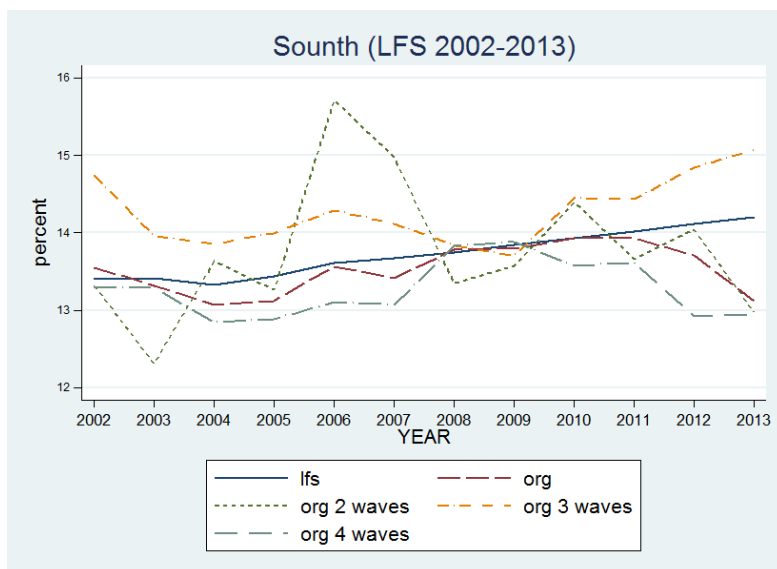
รูปที่ 78: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่อยู่ภาคเหนือ

จากรูปที่ 78 แสดงร้อยละกลุ่มตัวอย่างที่อยู่ภาคเหนือ พบว่า ข้อมูล LFS ร้อยละของคนที่อยู่ในภาคเหนือมีแนวโน้มคงที่อย่างต่อเนื่อง ในขณะที่ข้อมูลตัวอย่างซ้ำ ORG กลับมีแนวโน้มที่เพิ่มขึ้นอย่างช้า ๆ ซึ่งหากจะเปรียบเทียบร้อยละของคนที่อยู่ในภาคเหนือที่มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด คือข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 3 ครั้ง



รูปที่ 79: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่อยู่ภาคตะวันออกเฉียงเหนือ

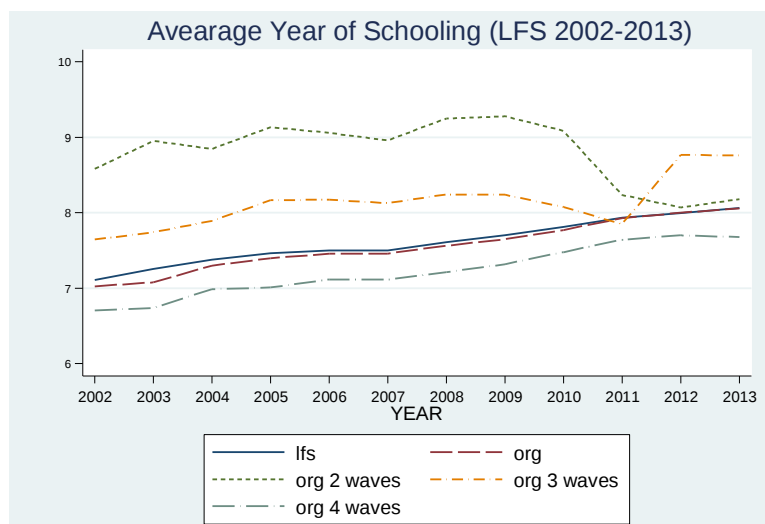
จากรูปที่ 79 แสดงร้อยละกลุ่มตัวอย่างที่อยู่ภาคตะวันออกเฉียงเหนือ พบว่า ทั้งข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีร้อยละของคนที่อยู่ในภาคตะวันออกเฉียงเหนือมีแนวโน้มคงที่อย่างต่อเนื่อง ซึ่งหากจะเปรียบเทียบร้อยละของคนที่อยู่ในภาคตะวันออกเฉียงเหนือที่มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด คือ ข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 4 ครั้ง



รูปที่ 80: ร้อยละของจำนวนกลุ่มตัวอย่างจากข้อมูล LFS และ ORG ที่อยู่ภาคใต้

จากรูปที่ 80 แสดงร้อยละกลุ่มตัวอย่างที่อยู่ภาคใต้ พบว่า ทั้งข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีร้อยละของคนที่อยู่ในภาคใต้เพิ่มขึ้นอย่างช้า ๆ ซึ่งหากจะเปรียบเทียบร้อยละของคนที่อยู่ในภาคใต้ที่มีค่าใกล้เคียงกับข้อมูล LFS มากที่สุด คือข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 4 ครั้ง

และเมื่อพิจารณาจำนวนปีการศึกษาของข้อมูล LFS และ ORG จากรูปที่ 81 แสดงจำนวนปีการศึกษาเฉลี่ยรายปี พบว่า ทั้งข้อมูล LFS และข้อมูลตัวอย่างซ้ำ ORG มีแนวโน้มเพิ่มขึ้นอย่างต่อเนื่อง โดยมีจำนวนปีการศึกษาเฉลี่ยอยู่ในช่วง 7-9 ปี ซึ่งเทียบเท่ากับระดับมัธยมศึกษาตอนต้น นั่นแสดงว่าประชากรไทยส่วนใหญ่ตั้งแต่ปี 2002-2013 จบการศึกษาในระดับชั้นมัธยมศึกษาตอนต้น แต่เมื่อเปรียบเทียบจำนวนปีการศึกษาเฉลี่ยของแต่ละชุดข้อมูล พบว่า ข้อมูล ORG ที่ถูกสัมภาษณ์ 2 ครั้ง มีจำนวนปีการศึกษาเฉลี่ยสูงกว่าข้อมูล LFS และข้อมูล ORG อื่นๆ โดยข้อมูลที่มีจำนวนปีการศึกษาเฉลี่ยใกล้เคียงกับข้อมูล LFS มากที่สุด คือข้อมูล ORG กลุ่มตัวอย่างซ้ำที่ถูกสัมภาษณ์ 4 ครั้ง



รูปที่ 81: จำนวนปีการศึกษาเฉลี่ยรายปีจากข้อมูล LFS และ ORG

ภาคผนวก ข.

ชุดคำสั่งในการจัดการข้อมูลรายบุคคลแบบตัวอย่างซ้ำจากข้อมูลภาวะการทำงาน
ของประชากร

คำอธิบาย
ชุดคำสั่งสำหรับการจัดการข้อมูลรายบุคคลแบบตัวอย่างซ้ำจากข้อมูลภาวการณ์ทำงาน
ของประชากร

1. ตั้งค่าโฟลเดอร์ต่าง ๆ เพื่อใช้ในการเก็บข้อมูลโดยมีโฟลเดอร์หลักชื่อ ORG Data และมีโฟลเดอร์ย่อยดังนี้

- raw data
- prepare data
- combine data
- keep data
- org data

```
***** RIPED VERSION DATE 3-11-2017 *****

clear
clear matrix
clear mata
set more off
set mem lg

//.....Set Path (Automatic: Do not change).....//
local fullpath c(pwd) /* display "`fullpath'" */
local splitat = strpos(`fullpath',"\ORG Data") /* display "`splitat'" */
global path = substr(`fullpath',1,`splitat') /* display "%{path}" */
sysdir set PERSONAL "%{path}ORG Data"

//....Set Path for this folder (Automatic: Do not change)....//

global raw "%{path}ORG Data\raw data/"
global pre "%{path}ORG Data\prepare data/"
global comb "%{path}ORG Data\combine data/"
global keep "%{path}ORG Data\keep data/"
global org "%{path}ORG Data\org data/"

*****
```

2. ตัดข้อมูลจากไฟล์ .dat เป็นไฟล์ stata โดยตั้งชื่อตัวแปรตามด้วยตำแหน่งของคอลัมน์อ้างอิงตาม datadict จากข้อมูล LFS ในแต่ละปีและไตรมาสที่ได้มาจากสำนักงานสถิติแห่งชาติ

```

**** (0) CUT RAW DATA IN 2002-2003 FOR SAMSET A ****

**>>>..... LFS 2002 Quarter 1 .....<<<
//***** (1) cut stata file from law data*****//
clear all
set more off
infix reg 1 ///
cwt 2-3 ///
area 4 ///
psu_no 5-8 ///
month 9-10 ///
yr 11-12 ///
hh_no 13-14 ///
hhtype 15 ///
members 16-17 ///
listing 18-19 ///
enum 20-21 ///
no 22-23 ///
relation 24 ///
sex 25 ///
age 26-27 ///
marital 28 ///
studynow 29 ///
edulevel 30-31 ///
wk_7day 32 ///
receive 33 ///
return 34 ///
absent 35 ///
seeking 36 ///
method 37 ///
available 38 ///
re_unavail 39 ///
re_no_seek 40 ///
dr_seek 41 ///
ever_wk 42 ///
re_unem 43 ///
dr_unem 44 ///
str_occup 45-48 ///
indus 49-52 ///
status 53 ///
size 54 ///
main_hr 55-56 ///
other_hr 57-58 ///
total_hr 59-60 ///
more_wk 61 ///
more_hr 62-63 ///
finding 64 ///
re_nomore 65 ///
wage_type 66 ///
amount 67-71 ///
approx 72-76 ///
bonus 77-81 ///
ot 82-86 ///
oth_money 87-91 ///
food 92-96 ///
cloth 97-101 ///
house 102-106 ///
oth_thing 107-111 ///
lst_men 112 ///
rotatgrp 113 ///
str_samset 114 ///
re_wk 130-131 ///
re_ed 132-133 ///
weight 136-144 using "${raw}LFS-2002\LFS145.dat"
gen year=2002
gen round=1
replace weight=weight/10000
drop yr

```

3. ระบุคำอธิบายของแต่ละตัวแปร

```
//***** (2) label variables*****//
lab var year "YEAR"
ren reg region
lab var region "Region (Corrected flipped code, 1985-recent)"
ren area areatype
lab var area "Area Type (Corrected code, 1985-recent)"
ren cwt province
lab var province "PROVINCE (Corrected for the flipped code 1985-recent)"
ren psu_no psunumbr
lab var psunumbr "PSU NUMBER (2002-2013)"
lab var samset "SAMPLING HOUSEHOLD (2002-recent)"
lab var rotatgrp "ROTATION GROUP (2002-2011)"
lab var month "Month (2001-recent)"
ren hh_no hhnumber
lab var hhnumber "HOUSEHOLD NUMBER (1985-recent)"
ren no memnumbr
lab var memnumbr "HOUSEHOLD MEMBER NUMBER (1985-recent)"
lab var hhtype "HOUSEHOLD TYPE (1985-recent)"
ren listing memlist
lab var memlist "NUMBER OF PERSONS IN HOUSEHOLD IN THE LISTING STEP (2001-recent)"
ren enum enumtype
lab var enumtype "THE HOUSEHOLD ENUMERATION SAMPLE (1985-recent)"
lab var relation "RELATIONSHIP TO HEAD OF HOUSEHOLD (1985-recent)"
ren members memenum
lab var memenum "THE NUMBER OF HOUSEHOLD MEMBER"
lab var sex "SEX (1985-recent)"
lab var age "AGE (1985-recent)"
lab var marital "MARITAL STATUS (1985-recent)"
lab var studynow "ATTENDING SCHOOL (1985-2006, see GRADE_A variable after 2006)"
lab var edulevel "Highest Education Level (1985-2006, see GRADE_B variable after 2006)"
ren wk_7day wrklstwk
lab var wrklstwk "WORK DURING THE SURVEY WEEK (1985-recent)"
lab var receive "Not working but still receive wages, salary or profit (2001-recent)"
ren return returnwk
lab var returnwk "RETURN TO WORK (2001-recent)"
lab var absent "How long you are not working (2001-recent)"
ren seeking seekjob1
lab var seekjob1 "LOOKING FOR A JOB (2001_Q3-recent)"
ren method howseek
lab var howseek "HOW TO LOOKING FOR A JOB (2001-recent)"
lab var available "Available for work last week 2013"
lab var re_unavail "THE REASON OF UNAVAILABLE WORK DURING THE SURVEY (Corrected code, 1985-recent)"
```

```

ren re_no_seek ynotseek
lab var ynotseek "THE REASON OF NOT LOOKING FOR A JOB IN THE SURVEY WEEK (1985-recent)"
ren dr_seek dr_seek
lab var dr_seek "THE NUMBER OF UNAVAILABLE WORKING DAY IN THE SURVEY WEEK (1985-2013)"
ren ever_wk everwork
lab var everwork "EVER WORK BEFORE (1985-recent)"
ren re_unem yquitjob
lab var yquitjob "THE REASON TO LEAVE THE LAST JOB (2001-recent)"
ren dr_unem dr_unem
lab var dr_unem "How long did you leave from the last job (2001-recent)"
ren occup occuptn
lab var occuptn "OCCUPATION (1985-recent)"
ren indus industry
lab var industry "INDUSTRY (1985-recent)"
ren status wkstatus_old2
lab var wkstatus_old2 "Work Status (2001-recent)"
ren size bizsize
lab var bizsize "Size of the Business (1998-recent)"
ren main_hr mainhrwk
lab var mainhrwk "THE NUMBER OF PRINCIPLE OCCUPATION HOURS (1985-recent)"
ren other_hr othhrwk
lab var othhrwk "THE NUMBER OF OTHER OCCUPATION HOURS (1985-recent)"
ren total_hr sumhrwk
lab var sumhrwk "TOTAL WORKING HOURS (1985-recent)"
ren more_wk wntwkmor
lab var wntwkmor "HAVING MORE HOURS AND WANT TO WORK MORE IN THE SURVEY WEEK (1985-recent)"
ren more_hr more_hr
lab var more_hr "Additional hours per week for working (2001-recent)"
ren finding finding
lab var finding "LOOKING FOR AN ADDITIONAL WORK (2001-recent)"
ren re_nomore ynowkmor
lab var ynowkmor "THE REASON OF NOT LOOKING FOR AN ADDITIONAL WORK (2001-recent)"
ren wage_type wagetype
lab var wagetype "WAGE TYPE (1985-recent)"
ren amount wagerate
lab var wagerate "WAGE RATE (1985-recent)"
ren approx salary
lab var salary "SALARY (1985-recent)"
ren bonus bonus
lab var bonus "BONUS PER YEAR (1985-recent)"
ren ot overtime
lab var overtime "AVERAGE OVERTIME WAGE PER MONTH (1985-recent)"
ren oth_money othmoney
lab var othmoney "AVERAGE OTHER EARNING PER MONTH (1985-recent)"
ren food food
lab var food "FOOD BENEFIT BESIDE WAGE (1985-recent)"
ren cloth cloth
lab var cloth "CLOTH BENEFIT BESIDE WAGE (1985-recent)"
ren house house
lab var house "HOUSE BENEFIT BESIDE WAGE (1985-recent)"
ren oth_thing oththing
lab var oththing "OTHER BENEFIT THING (2000-recent)"
ren lst_mem lastmem
lab var lastmem "THE LAST MEMBER OF HOUSEHOLD (1985-recent)"
ren re_wk re_wk_new
lab var re_wk_new "WORK CHARACTERISTICS RECODE (code correction, 1985-recent)"
ren re_ed re_ed
lab var re_ed "Education Recode 2002-recent"
lab var weight "WEIGHT (1985-recent)"
lab var round "ROUND/QUARTER (1985-recent)"

```

4. เปลี่ยนตัวแปรระดับการศึกษาในปี 2002-2006 ให้เป็นโค้ดเดียวกันกับปี 2007-2013

```
**replace re_ed
replace re_ed=16 if re_ed==14 & year<2007
replace re_ed=. if re_ed==15 & year<2007
replace re_ed=14 if (edulevel==32 | edulevel==69 | edulevel==83) & year<2007
replace re_ed=15 if (edulevel==33 | edulevel==70) & year<2007
replace re_ed=17 if (edulevel==96 | edulevel==98 | edulevel==99) & year<2007
```

5. สร้างตัวแปร ID รายบุคคลโดยตั้งชื่อว่า caseid (ทำขั้นตอนที่ 2-5 กับทุกข้อมูลในแต่ละปีและแต่ละไตรมาส)

```
//***** (3) gen caseid*****//
gen samset_new=samset
gen ea_set=rotatgrp
gen blk_vil=.
gen enumdist=.
tostring region, gen (reg_c)
tostring blk_vil, gen (blk_vil_c)
tostring memnumbr, gen (memnumbr_c)
tostring round, gen (round_c)
tostring province, gen (province_c)
tostring areatype, gen (areatype_c)
tostring hhnumber, gen (hhnumber_c)
tostring year, gen (year_c)
tostring psunumbr, gen (psunumbr_c)
tostring enumdist, gen (enumdist_c)
tostring hhtype, gen (hhtype_c)
tostring ea_set, gen (ea_set_c)

replace reg_c="0" if region==.
replace ea_set_c="0" if ea_set==.

replace hhnumber_c="0"+hhnumber_c if hhnumber<10
replace hhnumber_c="00" if hhnumber==.

replace blk_vil_c="00"+blk_vil_c if blk_vil<10
replace blk_vil_c="0"+blk_vil_c if (blk_vil>9 & blk_vil<100)
replace blk_vil_c="000" if blk_vil==.

replace psunumbr_c="000"+psunumbr_c if psunumbr<10
replace psunumbr_c="00"+psunumbr_c if psunumbr>9 & psunumbr<100
replace psunumbr_c="0"+psunumbr_c if psunumbr>99 & psunumbr<1000
replace psunumbr_c="0000" if psunumbr==.

replace enumdist_c="00"+enumdist_c if enumdist<10
replace enumdist_c="0"+enumdist_c if enumdist>9 & enumdist<100
replace enumdist_c="000" if enumdist==.

replace memnumbr_c="00" if memnumbr==.
replace memnumbr_c="0"+ memnumbr_c if memnumbr<10

gen caseid = reg_c + province_c + areatype_c + psunumbr_c + samset_new + ea_set_c + ///
enumdist_c + blk_vil_c + hhtype_c + hhnumber_c + memnumbr_c

drop reg_c areatype_c round_c year_c hhnumber_c hhtype_c memnumbr_c province_c psunumbr_c ///
enumdist_c blk_vil_c hhnumber_c memnumbr_c ea_set_c samset_new

duplicates list caseid

lab var caseid "CASE IDENTIFICATION NO. - NSO"
lab var ea_set "(2009, 2010, 2012, 2013)"
lab var blk_vil "Block/Village (1985-recent, except 2002 & 2003)"
lab var enumdist "ENUMERATION DISTRICT NUMBER (2004-recent)"
save "{pre}lfs2002q1.dta", replace
```

6. รวมข้อมูลในแต่ละ rotation plan ซึ่งตัวอย่างโค้ดคำสั่งตัวเป็นการรวมข้อมูล LFS ปี 2002-2003 สำหรับ samset A และสร้างตัวแปร ID ครั้วเรือนที่เกิดจากการตัด ID รายบุคคลใน 2 ตำแหน่งสุดท้ายทิ้ง และตั้งชื่อว่า casehid นอกจากนี้ยังสร้างตัวแปรชื่อ dum_enum เพื่อระบุประเภทของครั้วเรือนที่เป็นครั้วเรือนทดแทน โดยในการค้นหาครั้วเรือนตัวอย่างซ้ำ นักวิจัยจะพิจารณาเฉพาะครั้วเรือนที่ยังคงใช้หมายเลขเดิมและไม่เป็นครั้วเรือนทดแทน

```

****# (1) COMBINE DATASET 2002-2003 FOR SAMSET A ****#
*2002
use "{pre}lfs2002q1.dta",clear
append using "{pre}lfs2002q2.dta"
append using "{pre}lfs2002q3.dta"
append using "{pre}lfs2002q4.dta"
*2003
append using "{pre}lfs2003q1.dta"
append using "{pre}lfs2003q2.dta"
append using "{pre}lfs2003q3.dta"
append using "{pre}lfs2003q4.dta"

** generate casehid
gen casehid=substr(caseid,1,19)

** replace missing in rotatgrp in 2009,2010,2012,2013
replace rotatgrp=ea_set if rotatgrp==.

** generate enumtype in first and last of rotation group
sort casehid samset rotatgrp year round
bysort casehid samset rotatgrp : gen firstenum=enumtype[1]
bysort casehid samset rotatgrp : gen lastenum=enumtype[_N]

** count household by enumtype
gen dum_enum=.
replace dum_enum=1 if firstenum<lastenum & lastenum<=16
replace dum_enum=2 if firstenum>lastenum & lastenum<=16
replace dum_enum=0 if dum_enum==.
by casehid,sort : gen numhh=1 if _n==1
lab define dum_enum 0 "the same household" 1 "replacement household" 2 "replacement household become sample"
lab value dum_enum dum_enum

** label variables
lab var casehid "CASE IDENTIFICATION OF HOUSEHOLD"
lab var firstenum "checking household type in first wave"
lab var lastenum "checking household type in last wave"
lab var dum_enum "dummy of household type"
lab var numhh "count unique household number"

save "{comb}lfs2002-2003.dta",replace
*****

```

7. เก็บข้อมูลแยกตาม rotation plan และสร้างตัวแปร wave หรือจำนวนครั้งที่ถูกสัมภาษณ์ซึ่งจะมีได้สูงสุดคือ 4 ครั้ง เพื่อใช้ตรวจสอบและค้นหาข้อมูลตัวอย่างซ้ำรายบุคคล
- จัดเก็บข้อมูลสำหรับ samset A และ rotation 1

```

**** (2) SAVE AS DATA FOREACH ROTATION IN SAMSET A ****

** 1 save sub-panel ORG samset=A,rotation=1 **
use "{comb}lfs2002-2003.dta",clear
keep if samset=="A" & rotatgrp==1 & dum_enum==0
sort casehid year round caseid
order casehid caseid round year sex age relation marital edulevel memenum enumtype

** generate wave in samaset A rotation 1
gen wave=1 if year==2002 & round==1
replace wave=2 if year==2002 & round==4
replace wave=3 if year==2003 & round==1
replace wave=4 if year==2003 & round==4
order casehid caseid wave round year sex age relation marital edulevel memenum enumtype
sort caseid wave

** check other round in ORG plan **
drop if wave==.

**generate wave dummies
tab wave,gen(wave_)

**fill max value of wave dummy by household and individual
forv i=1/4 {
  bysort caseid : egen waveid_`i' =max(wave_`i')
  bysort casehid : egen wavehid_`i' =max(wave_`i')
}

**drop variables we don't need
drop wave_*
order casehid caseid wave round year sex age relation marital edulevel memenum enumtype waveid_* wavehid_*

**count number of wave by hid
egen numwave=rowtotal(wavehid_1 wavehid_2 wavehid_3 wavehid_4)

forv i=1/4 {
  **count household in `i' waves
  by casehid,sort : gen no_wave`i'=1 if _n==1 & numwave==`i'
  replace no_wave`i'=0 if no_wave`i'==.
}

**gen number of total household
by casehid,sort : gen total_hhnum=1 if _n==1
sort casehid year round caseid
count if total_hhnum==1

**gen number of total members
by caseid,sort : gen total_memnum=1 if _n==1
sort casehid year round caseid
count if total_memnum==1

** generate start and end in wavehid
forv i=1/4 {
  by casehid,sort :gen w`i'=`i' if _n==1 & wavehid_`i'==1
  replace w`i'=0 if w`i'==.
}

**combine wave of hid
gen wavehid=w1*1000+w2*100+w3*10+w4

```

```

**change to correct wavehid
replace wavehid=4 if wavehid==4
replace wavehid=3 if wavehid==30
replace wavehid=34 if wavehid==34
replace wavehid=2 if wavehid==200
replace wavehid=24 if wavehid==204
replace wavehid=23 if wavehid==230
replace wavehid=234 if wavehid==234
replace wavehid=1 if wavehid==1000
replace wavehid=14 if wavehid==1004
replace wavehid=13 if wavehid==1030
replace wavehid=134 if wavehid==1034
replace wavehid=12 if wavehid==1200
replace wavehid=124 if wavehid==1204
replace wavehid=123 if wavehid==1230
replace wavehid=1234 if wavehid==1234

**sort and reorder data
sort casehid year round caseid
drop w1 w2 w3 w4 wavehid *
order casehid caseid wave round year sex age relation marital edulevel memenum enumtype waveid_1 numwave wavehid

** generate start and end in waveid
forv i=1/4 {
  by caseid, sort :gen w`i'=`i' if waveid_1=`i'==1
  replace w`i'=0 if w`i'==.
}

**combine wave of id
gen waveid=w1*1000+w2*100+w3*10+w4

**change to correct waveid
replace waveid=3 if waveid==30
replace waveid=2 if waveid==200
replace waveid=24 if waveid==204
replace waveid=23 if waveid==230
replace waveid=234 if waveid==234
replace waveid=1 if waveid==1000
replace waveid=14 if waveid==1004
replace waveid=13 if waveid==1030
replace waveid=134 if waveid==1034
replace waveid=12 if waveid==1200
replace waveid=124 if waveid==1204
replace waveid=123 if waveid==1230
replace waveid=1234 if waveid==1234

**sort and reorder data
sort casehid year round caseid
drop w1 w2 w3 w4 waveid_1 waveid_2 waveid_3 waveid_4
order casehid caseid wave wavehid waveid
sort casehid caseid wave

**label variables didn't have description
lab var wave "wave"
lab var wavehid "checking wave of households"
lab var waveid "checking wave of individuals"
lab var numwave "number of household's wave"
drop no_wave*
lab var total_hhnum "marking for unique household (casehid)"
lab var total_memnum "marking for unique individuals (caseid)"

save "${keep}samsetA_rotation1.dta", replace
*-----*

```

- จัดเก็บข้อมูลสำหรับ samset A และ rotation 2

```

** 2 save sub-panel ORG samset=A,rotation=2 **
use "{comb}lfs2002-2003.dta",clear
keep if samset=="A" & rotatgrp==2 & dum_enum==0
sort casehid year round caseid
order caseid round year sex age relation marital edulevel memenum enumtype

**generate wave in samset=A,rotation=2
gen wave=1 if year==2002 & round==1
replace wave=2 if year==2002 & round==2
replace wave=3 if year==2003 & round==1
replace wave=4 if year==2003 & round==2
order casehid caseid wave round year sex age relation marital edulevel memenum enumtype
sort caseid wave

** check other round in ORG plan **
drop if wave==.

**generate wave
tab wave,gen(wave_)
forv i=1/4 {
  bysort caseid : egen waveid_`i' =max(wave_`i')
  bysort casehid : egen wavehid_`i' =max(wave_`i')
}
drop wave_*
order casehid caseid wave round year sex age relation marital edulevel memenum enumtype waveid_* wavehid_*

**count number of wave by hid
egen numwave=rowtotal(wavehid_1 wavehid_2 wavehid_3 wavehid_4)

**count household in k waves
forvalue k = 1/4 {
  by casehid,sort : gen no_wave`k'=1 if _n==1 & numwave==`k'
  replace no_wave`k'=0 if no_wave`k'==.
}

**gen number of total household
by casehid,sort : gen total_hhnum=1 if _n==1
sort casehid year round caseid

**gen number of total members
by caseid,sort : gen total_memnum=1 if _n==1
sort casehid year round caseid
count if total_memnum==1

** generate start and end in wavehid
forv i=1/4 {
  by casehid,sort :gen w`i'=`i' if _n==1 & wavehid_`i'==1
  replace w`i'=0 if w`i'==.
}

**combine wave of hid
gen wavehid=w1*1000+w2*100+w3*10+w4

**change to correct wavehid
replace wavehid=4 if wavehid==4
replace wavehid=3 if wavehid==30
replace wavehid=34 if wavehid==34
replace wavehid=2 if wavehid==200
replace wavehid=24 if wavehid==204
replace wavehid=23 if wavehid==230
replace wavehid=234 if wavehid==234
replace wavehid=1 if wavehid==1000
replace wavehid=14 if wavehid==1004
replace wavehid=13 if wavehid==1030
replace wavehid=134 if wavehid==1034
replace wavehid=12 if wavehid==1200
replace wavehid=124 if wavehid==1204
replace wavehid=123 if wavehid==1230
replace wavehid=1234 if wavehid==1234

```

```

sort casehid year round caseid
drop w1 w2 w3 w4 wavehid_*
order casehid caseid wave round year sex age relation marital edulevel memenum enumtype waveid_* numwave wavehid

** generate start and end in waveid
forv i=1/4 {
  by caseid,sort :gen w`i'=`i' if waveid_`i'==1
  replace w`i'=0 if w`i'==.
}

**combine wave of id
gen waveid=w1*1000+w2*100+w3*10+w4

**change to correct waveid
replace waveid=3 if waveid==30
replace waveid=2 if waveid==200
replace waveid=24 if waveid==204
replace waveid=23 if waveid==230
replace waveid=234 if waveid==234
replace waveid=1 if waveid==1000
replace waveid=14 if waveid==1004
replace waveid=13 if waveid==1030
replace waveid=134 if waveid==1034
replace waveid=12 if waveid==1200
replace waveid=124 if waveid==1204
replace waveid=123 if waveid==1230
replace waveid=1234 if waveid==1234

**sort and reorder data
sort casehid year round caseid
drop w1 w2 w3 w4 waveid_1 waveid_2 waveid_3 waveid_4
order casehid caseid wave wavehid waveid
sort casehid caseid wave

**label variables didn't have descreption
lab var wave "wave"
lab var wavehid "checking wave of households"
lab var waveid "checking wave of individuals"
lab var numwave "number of household's wave"
drop no_wave*
lab var total_hhnum "marking for unique household (casehid)"
lab var total_memnum "marking for unique individuals (caseid)"

save "${keep}samsetA_rotation2.dta",replace
*-----*

```

- จัดเก็บข้อมูลสำหรับ samset A และ rotation 3

```

** 3 save sub-panel ORG samset=A,rotation=3 **
use "{comb}ifs2002-2003.dta",clear
keep if samset=="A" & rotatgrp==3 & dum_enum==0
sort casehid year round caseid
order casehid caseid round year sex age relation marital edulevel memenum enumtype

gen wave=1 if year==2002 & round==2
replace wave=2 if year==2002 & round==3
replace wave=3 if year==2003 & round==2
replace wave=4 if year==2003 & round==3

order casehid caseid wave round year sex age relation marital edulevel memenum enumtype
sort caseid wave

** check other round in ORG plan **
drop if wave==.

**generate wave
tab wave,gen(wave_)
forv i=1/4 {
  bysort caseid : egen waveid_`i' =max(wave_`i')
  bysort casehid : egen wavehid_`i' =max(wave_`i')
}
drop wave_*
order casehid caseid wave round year sex age relation marital edulevel memenum enumtype waveid_* wavehid_*
egen numwave=rowtotal(wavehid_1 wavehid_2 wavehid_3 wavehid_4)

**count household in k waves
forvalue k = 1/4 {
  by casehid,sort : gen no_wave`k'=1 if _n==1 & numwave==`k'
  replace no_wave`k'=0 if no_wave`k'==.
}

**gen number of total household
by casehid,sort : gen total_hhnum=1 if _n==1
sort casehid year round caseid

**gen number of total members
by caseid,sort : gen total_memnum=1 if _n==1
sort casehid year round caseid
count if total_memnum==1

** generate start and end in wavehid
forv i=1/4 {
  by casehid,sort :gen w`i'=`i' if _n==1 & wavehid_`i'==1
  replace w`i'=0 if w`i'==.
}

**combine wave of hid
gen wavehid=w1*1000+w2*100+w3*10+w4

**change to correct wavehid
replace wavehid=4 if wavehid==4
replace wavehid=3 if wavehid==30
replace wavehid=34 if wavehid==34
replace wavehid=2 if wavehid==200
replace wavehid=24 if wavehid==204
replace wavehid=23 if wavehid==230
replace wavehid=234 if wavehid==234
replace wavehid=1 if wavehid==1000
replace wavehid=14 if wavehid==1004
replace wavehid=13 if wavehid==1030
replace wavehid=134 if wavehid==1034
replace wavehid=12 if wavehid==1200
replace wavehid=124 if wavehid==1204
replace wavehid=123 if wavehid==1230
replace wavehid=1234 if wavehid==1234

```

```

**sort and reorder data
sort casehid year round caseid
drop w1 w2 w3 w4 wavehid_*
order casehid caseid wave round year sex age relation marital edulevel memenum enumtype waveid_* numwave wavehid

** generate start and end in waveid
forv i=1/4 {
  by caseid,sort :gen w`i'=`i' if waveid_`i'==1
  replace w`i'=0 if w`i'==.
}

**combine wave of id
gen waveid=w1*1000+w2*100+w3*10+w4

**change to correct waveid
replace waveid=3 if waveid==30
replace waveid=2 if waveid==200
replace waveid=24 if waveid==204
replace waveid=23 if waveid==230
replace waveid=234 if waveid==234
replace waveid=1 if waveid==1000
replace waveid=14 if waveid==1004
replace waveid=13 if waveid==1030
replace waveid=134 if waveid==1034
replace waveid=12 if waveid==1200
replace waveid=124 if waveid==1204
replace waveid=123 if waveid==1230
replace waveid=1234 if waveid==1234

**sort and reorder data
sort casehid year round caseid
drop w1 w2 w3 w4 waveid_1 waveid_2 waveid_3 waveid_4
order casehid caseid wave wavehid waveid
sort casehid caseid wave

**label variables didn't have descreption
lab var wave "wave"
lab var wavehid "checking wave of households"
lab var waveid "checking wave of individuals"
lab var numwave "number of household's wave"
drop no_wave*
lab var total_hhnum "marking for unique household (casehid)"
lab var total_memnum "marking for unique individuals (caseid)"

save "${keep}samsetA_rotation3.dta",replace
*-----*

```

- จัดเก็บข้อมูลสำหรับ samset A และ rotation 4

```

** 4 save sub-panel ORG samset=A,rotation=4 **
use "{comb}lfs2002-2003.dta",clear
keep if samset=="A" & rotatgrp==4 & dum_enum==0
sort casehid year round caseid
order casehid caseid round year sex age relation marital edulevel memenum enumtype

gen wave=1 if year==2002 & round==3
replace wave=2 if year==2002 & round==4
replace wave=3 if year==2003 & round==3
replace wave=4 if year==2003 & round==4

order casehid caseid wave round year sex age relation marital edulevel memenum enumtype
sort caseid wave

** check other round in ORG plan **
drop if wave==.

**generate wave
tab wave,gen(wave_)
forv i=1/4 {
  bysort caseid : egen waveid_`i' =max(wave_`i')
  bysort casehid : egen wavehid_`i' =max(wave_`i')
}
drop wave_*
order casehid caseid wave round year sex age relation marital edulevel memenum enumtype waveid_* wavehid_*
egen numwave=rowtotal(wavehid_1 wavehid_2 wavehid_3 wavehid_4)

**count household in k waves
forvalue k = 1/4 {
  by casehid,sort : gen no_wave`k'=1 if _n==1 & numwave==`k'
  replace no_wave`k'=0 if no_wave`k'==.
}

```

```

**gen number of total household
by casehid,sort : gen total_hhnum=1 if _n==1
sort casehid year round caseid

**gen number of total members
by caseid,sort : gen total_memnum=1 if _n==1
sort casehid year round caseid
count if total_memnum==1

** generate start and end in wavehid
forv i=1/4 {
by casehid,sort :gen w`i'=`i' if _n==1 & wavehid_`i'==1
replace w`i'=0 if w`i'==.
}

**combine wave of hid
gen wavehid=w1*1000+w2*100+w3*10+w4

**change to correct wavehid
replace wavehid=4 if wavehid==4
replace wavehid=3 if wavehid==30
replace wavehid=34 if wavehid==34
replace wavehid=2 if wavehid==200
replace wavehid=24 if wavehid==204
replace wavehid=23 if wavehid==230
replace wavehid=234 if wavehid==234
replace wavehid=1 if wavehid==1000
replace wavehid=14 if wavehid==1004
replace wavehid=13 if wavehid==1030
replace wavehid=134 if wavehid==1034
replace wavehid=12 if wavehid==1200
replace wavehid=124 if wavehid==1204
replace wavehid=123 if wavehid==1230
replace wavehid=1234 if wavehid==1234

sort casehid year round caseid
drop w1 w2 w3 w4 wavehid *
order casehid caseid wave round year sex age relation marital edulevel memenum enumtype waveid_* numwave wavehid

** generate start and end in waveid
forv i=1/4 {
by caseid,sort :gen w`i'=`i' if waveid_`i'==1
replace w`i'=0 if w`i'==.
}

**combine wave of id
gen waveid=w1*1000+w2*100+w3*10+w4

**change to correct waveid
replace waveid=3 if waveid==30
replace waveid=2 if waveid==200
replace waveid=24 if waveid==204
replace waveid=23 if waveid==230
replace waveid=234 if waveid==234
replace waveid=1 if waveid==1000
replace waveid=14 if waveid==1004
replace waveid=13 if waveid==1030
replace waveid=134 if waveid==1034
replace waveid=12 if waveid==1200
replace waveid=124 if waveid==1204
replace waveid=123 if waveid==1230
replace waveid=1234 if waveid==1234

**sort and reorder data
sort casehid year round caseid
drop w1 w2 w3 w4 waveid_1 waveid_2 waveid_3 waveid_4
order casehid caseid wave wavehid waveid
sort casehid caseid wave

**label variables didn't have descreption
lab var wave "wave"
lab var wavehid "checking wave of households"
lab var waveid "checking wave of individuals"
lab var numwave "number of household's wave"
drop no_wave*
lab var total_hhnum "marking for unique household (casehid)"
lab var total_memnum "marking for unique individuals (caseid)"

save "${keep}samsetA_rotation4.dta",replace
*****

```

8. สร้างเงื่อนไขเพื่อตรวจสอบตัวอย่างซ้ำรายบุคคลในแต่ละชุดข้อมูลที่แยกเก็บตาม rotation plan โดยใช้ caseid ในการสร้างเงื่อนไข

```
*****
#### .....
#### .....
#### .....
#### .....
#### .....
#### ..... ID MATCHED DATA.....
#### .....
#### .....
#### .....
#### .....
#### .....
#### .....
#### .....
#### .....
#### .....
#### .....
#### .....
#### .....
*****
```

- สร้างตัวแปร wavehid เพื่อระบุว่าแต่ละครัวเรือนถูกสัมภาษณ์ใน wave ไດบ้าง และสร้างตัวแปร dummy ที่มีชื่อ samenumb เพื่อตรวจสอบว่าในแต่ละ wave มีสมาชิกครัวเรือนถูกสัมภาษณ์เท่ากันหรือไม่

```
**** (3) SAVE SUB-PANEL ORG DATA FOR SAMSET A ROTATION 1-4 ****
set more off
forvalue i=1/4 {
    use "${keep}samsetA_rotation`i'.dta", clear

    **(1) filling wave in the same household **
    bysort casehid : egen wavehid2=max(wavehid)
    order wavehid2,after(wavehid)
    lab var wavehid2 "filling wave in the same household"
    *-----*

    ** (2) count number of membes **
    bysort casehid wave : gen numb_members=_N
    bysort casehid : egen max_numb=max(numb_members)
    bysort casehid : egen min_numb=min(numb_members)
    lab var numb_members "counting on number of household member by wave"
    lab var max_numb "counting on maximum number of household member by wave"
    lab var min_numb "counting on minimum number of household member by wave"
    ** gen dummy samenumb
    gen samenumb=1 if max_numb!=min_numb
    replace samenumb=0 if max_numb==min_numb
    lab def samenumb 0 "same" 1 "not same"
    lab value samenumb samenumb
    lab var samenumb "the same number of household member dummy"
    *-----*
```

- สร้างเงื่อนไขเพื่อตรวจสอบข้อมูลตัวอย่างซ้ำรายบุคคล โดยใช้ caseid เมื่อกำหนดให้คนเดียวกันต้องมีเพศเดียวกัน และมีอายุห่างกันไม่เกิน 2 ปีนับจากการถูกสัมภาษณ์ในครั้งแรกจนถึงครั้งสุดท้าย และตั้งชื่อตัวแปรที่สอดคล้องกับเงื่อนไขนี้ว่า same_member1

```

***** STEP 1 check same member if age>=0 & age<=2 *****

** (3) check same member **
** 3.1 - gen mode age
bysort caseid sex : egen modeage=mode(age)
sort casehid caseid wave
gen diffmode=abs(modeage-age)
lab var modeage "median of age by caseid"
lab var diffmode "absolute difference value between median age and age"

** 3.2 - gen mean age
bysort caseid sex: egen meanage=mean(age)
sort casehid caseid wave
gen f_meanage=floor(meanage)
gen diffmean=abs(f_meanage-age)
lab var meanage "mean of age by caseid"
lab var diffmean "absolute difference value between mean age and age"

** 3.3 - combine difference modeage and meanage
gen diffage=diffmode
replace diffage=diffmean if diffage==.
replace diffage=. if diffage>2
*drop
drop modeage diffmode meanage f_meanage diffmean
lab var diffage "absolute difference value between expected age and real age"

** 3.4 - gen mode sex
bysort caseid : egen modesex=mode(sex) if diffage!=.
sort casehid caseid wave
lab var modesex "median of sex by caseid"
**add
bysort caseid : egen meansex=mean(sex)
replace meansex=floor(meansex)
replace modesex=meansex if modesex==.
drop meansex

** 3.5 - gen age in first and last wave by caseid
* first age
egen group=group(caseid) if modesex!=.
bysort group : egen wavel=min(wave) if group!=.
bysort group : gen age1=age if wave==wavel
bysort group : egen agefirst=max(age1)
sort casehid caseid wave group
drop wavel age1
lab var agefirst "first age in first wave"
*last age
bysort group : egen wave_last=max(wave) if group!=.
bysort group : gen age_last=age if wave==wave_last
bysort group : egen agelast=max(age_last)
drop wave_last age_last
lab var agelast "last age in last wave"

** 3.6 - difference value between age and first age and last age
*diff first age
gen diff_agefirst=age-agefirst
replace diff_agefirst=. if diff_agefirst>2 | diff_agefirst<-2
lab var diff_agefirst "difference value between first age and age"
*diff last age
gen diff_agelast=age-agelast
lab var diff_agelast "difference value between last age and age"
**add
replace diffage=diff_agefirst if diffage==.

```

```

** 3.7 - checking increased age by caseid
gen increase=1 if (diff_agefirst>0 & diff_agefirst!=.) & (diff_agelast<=0 & diff_agelast!=.)
replace increase=0 if increase==.
sort casehid caseid wave group
drop diff_agefirst agefirst diff_agelast agelast
lab var increase "increased age dummy"

** 3.8 - gen dummy for the same individuals
bysort caseid: gen same_member=1 if diffage>0 & diffage<=2 & sex==modesex & increase==1 & group!=.
replace same_member=0 if same_member==.
lab def same_member 1 "same" 0 "not same"
lab value same_member same_member
lab var same_member "the same individual dummy"
*-----*

** (4) re-checking the same individual live in more than one wave **

** 4.1 - gen dummy of waveid
gen dumwaveid=0 if waveid==1 | waveid==2 | waveid==3 | waveid==4
replace dumwaveid=1 if dumwaveid==.
lab def dumwaveid 0 "live in 1 wave" 1 "live in more than 1 wave"
lab value dumwaveid dumwaveid
lab var dumwaveid "individual's wave dummy"

** 4.2 - counting on individual panel **
gen count=1 if same_member==1
replace count=0 if count==.
bysort caseid: egen countw_samemem=sum(count)
lab var count "generated number for checking in the same individual"
lab var countw_samemem "count number of the same individual"

** 4.3 - gen same_member last
gen same_member_last=1 if same_member==1 & countw_samemem!=1
replace same_member_last=0 if same_member_last==.
lab def same_member_last 1 "same" 0 "not same"
lab value same_member_last same_member_last
lab var same_member_last "the same individual live in more than one wave"

** 4.4 - generate laour and household head dummy
*dummy of labour
gen labour=0 if age<15 | age>57
replace labour=1 if age>=15 & age<=57
lab def labour 0 "non labour" 1 "labour"
lab value labour labour

*dummy of household head change
bysort caseid : egen counthead=sum(relation) if relation==1
gen head0=1 if counthead==max_numb
replace head0=0 if counthead!=max_numb
bysort casehid : egen head=max(head0)
lab def head 0 "change" 1 "not change"
lab value head head

*label define marital
lab def marital 1 "Single" 2 "Married" 3 "Widowed" 4 "Divorced" 5 "Separated" 6 "Married but not known status"
lab value marital marital
*label define re_ed
lab def re_ed 1 "None" 2 "Less than elementary" 3 "Primary education" ///
4 "Lower secondary education" 5 "Upper Secondary-Academic" ///
6 "Upper Secondary-Vocational" 7 "Upper Secondary-Teacher Training" ///
8 "Post-secondary-Academic" 9 "Post-secondary-Higher Technical Education" ///
10 "Post-secondary-Teacher Training" 11 "Bachelor-Academic" ///
12 "Bachelor-Higher Technical Education" 13 "Bachelor-Teacher Training" ///
14 "Master degree level" 15 "Doctoral degee level" 16 "Otherse education" ///
17 "Unknown education"
lab value re_ed re_ed

drop counthead head0
lab var head "household head dummy"
lab var labour "labour dummy"

** save as data
save "${keep}samsetA_rotation'i'_keep.dta",replace
*-----*

```

```

** (5) keep only individuals live in more than one wave **
drop if same_member_last==0
by caseid, sort : gen count_memnum=1 if _n==1
sort casehid caseid wave
lab var count_memnum "count unique of individual live in more than one wave"
*****

```

- สร้างเงื่อนไขเพื่อตรวจสอบข้อมูลตัวอย่างซ้ำรายบุคคล โดยใช้ caseid เมื่อกำหนดให้คนเดียวกันต้องมีเพศเดียวกัน และมีอายุห่างกันไม่เกิน 1 ปีนับจากการถูกสัมภาษณ์ในครั้งแรกจนถึงครั้งสุดท้าย และตั้งชื่อตัวแปรที่สอดคล้องกับเงื่อนไขนี้ว่า same_member2

```

***** STEP 2 check same member if age>=0 & age<=1 *****

** (6) check same member **
drop diffage increase modesex count_memnum

** 6.1 - gen mode age
bysort caseid sex : egen modeage=mode(age)
sort casehid caseid wave
gen diffmode=abs(modeage-age)
lab var modeage "median of age by caseid"
lab var diffmode "absolute difference value between median age and age"

** 6.2 - gen mean age
bysort caseid sex: egen meanage=mean(age)
sort casehid caseid wave
gen f_meanage=floor(meanage)
gen diffmean=abs(f_meanage-age)
lab var meanage "mean of age by caseid"
lab var diffmean "absolute difference value between mean age and age"

** 6.3 - combine difference modeage and meanage
gen diffage=diffmode
replace diffage=diffmean if diffage==.
replace diffage=. if diffage>1
*drop
drop modeage diffmode meanage f_meanage diffmean
lab var diffage "absolute difference value between expected age and real age"

** 6.4 - gen mode sex
bysort caseid : egen modesex=mode(sex) if diffage!=.
sort casehid caseid wave
lab var modesex "median of sex by caseid"
**add
bysort caseid : egen meansex=mean(sex)
replace meansex=floor(meansex)
replace modesex=meansex if modesex==.
drop meansex

** 6.5 - gen age in first and last wave by caseid
* first age
bysort group : egen wavel=min(wave) if group!=.
bysort group : gen age1=age if wave==wavel
bysort group : egen agefirst=max(age1)
sort casehid caseid wave group
drop wavel age1
lab var agefirst "first age in first wave"
*last age
bysort group : egen wave_last=max(wave) if group!=.
bysort group : gen age_last=age if wave==wave_last
bysort group : egen agelast=max(age_last)
drop wave_last age_last
lab var agelast "last age in last wave"

```

```

** 6.6 - difference value between age and first age and last age
*diff first age
gen diff_agefirst=age-agefirst
replace diff_agefirst=. if diff_agefirst>1 | diff_agefirst<-1
lab var diff_agefirst "difference value between first age and age"
*diff last age
gen diff_agelast=age-agelast
lab var diff_agelast "difference value between last age and age"
**add
replace diffage=diff_agefirst if diffage==.

** 6.7 - checking increased age by caseid
gen increase=1 if (diff_agefirst>0 & diff_agefirst!=.) & (diff_agelast<0 & diff_agelast!=.)
replace increase=0 if increase==.
sort casehid caseid wave group
drop diff_agefirst agefirst diff_agelast agelast
lab var increase "increased age dummy"

** 6.8 - gen dummy for the same individuals
bysort caseid: gen same_member2=1 if diffage>0 & diffage<=1 & sex==modesex & increase==1 & group!=.
replace same_member2=0 if same_member2==.
**count person live in more than 1 wave
sort casehid caseid wave
bysort caseid: egen number_waveid=sum(count) if same_member2==1
**define only the same person live in more than 1 wave
replace same_member2=0 if number_waveid==1 & number_waveid!=.
lab def same_member2 1 "same" 0 "not same"
lab value same_member2 same_member2
lab var same_member2 "the same individual dummy step2 (age 0-1)"

** re-order variables
drop number_waveid
order count* total* samenumb*,after(increase )
order same_member same_member_last,after( samenumb )
ren same_member_last same_member1
lab var same_member1 "the same individual dummy step1 (age 0-2)"
*-----*

```

- สร้างเงื่อนไขเพื่อตรวจสอบข้อมูลตัวอย่างซ้ำรายบุคคล โดยใช้ caseid เมื่อกำหนดให้คนเดียวกันต้องมีเพศเดียวกัน และมีอายุห่างกันไม่เกิน 1 ปีนับจากการถูกสัมภาษณ์ในครั้งแรกจนถึงครั้งสุดท้าย นอกจากนี้ยังเพิ่มเงื่อนไขของสถานภาพสมรส และตั้งชื่อตัวแปรที่สอดคล้องกับเงื่อนไขว่า same_member3

```

***** STEP 3-4 check marital status and marital status change *****

** (7) check same member **
** 7.1 - generate marital change dummy **
*browse casehid caseid wave round year sex age marital same_member1 same_member2
bysort caseid : egen number_waveid0=sum(count) if same_member2==1
bysort caseid : egen number_waveid=max(number_waveid0)
drop number_waveid0
sort casehid caseid wave
bysort caseid marital: egen dum_marital0=sum(count) if same_member2==1
replace number_waveid=. if dum_marital0==.
gen dum_marital=number_waveid-dum_marital0
replace dum_marital=1 if dum_marital>0 & dum_marital!=.
lab def dum_marital 0 "not change" 1 "change"
lab value dum_marital dum_marital
lab var dum_marital "marital status dummy"
drop dum_marital0
*tab dum_marital if count_memnum==1

```

```

** 7.2 - checking marital status by wave **
sort casehid caseid wave
*wave1
bysort caseid : gen marital_0=marital if _n==1 & dum_marital!=.
bysort caseid : egen marital_1=max(marital_0)
drop marital_0
*wave2
bysort caseid : gen marital_0=marital if _n==2 & dum_marital!=.
bysort caseid : egen marital_2=max(marital_0)
drop marital_0
*wave3
bysort caseid : gen marital_0=marital if _n==3 & dum_marital!=.
bysort caseid : egen marital_3=max(marital_0)
drop marital_0
*wave4
bysort caseid : gen marital_0=marital if _n==4 & dum_marital!=.
bysort caseid : egen marital_4=max(marital_0)
drop marital_0

**replace dum_marital=not change if age wave1 < 15 (marital_1==.)
replace dum_marital=0 if marital_1==.

** 7.3 - generate same person with marital status condition **
tostring marital_1 marital_2 marital_3 marital_4,replace

gen str marital_combine=""
replace marital_combine=marital_1+marital_2+marital_3+marital_4
replace marital_combine = subinstr(marital_combine, ".", "",.)

gen marital_check=0

** number of wave = 2 **
*marital = 1 -> 1,2 (6)
replace marital_check=1 if marital_combine=="11" | marital_combine=="12" ///
| marital_combine=="16"
*marital = 2 -> 2(6),3,4,5
replace marital_check=1 if marital_combine=="22" | marital_combine=="26" ///
| marital_combine=="23" | marital_combine=="24" | marital_combine=="25"
*marital = 6 -> 6(2),3,4,5
replace marital_check=1 if marital_combine=="66" | marital_combine=="62" ///
| marital_combine=="63" | marital_combine=="64" | marital_combine=="65"
*marital = 3 -> 2(6),3
replace marital_check=1 if marital_combine=="32" | marital_combine=="36" ///
| marital_combine=="33"
*marital = 4 -> 2(6),3,4
replace marital_check=1 if marital_combine=="42" | marital_combine=="46" ///
| marital_combine=="43" | marital_combine=="44"
*marital = 5 -> 3,4,5
replace marital_check=1 if marital_combine=="53" | marital_combine=="54" ///
| marital_combine=="55"

** number of wave = 3 **
*marital = 1 -> 1 -> 1
replace marital_check=1 if marital_combine=="111"
*marital = 1 -> 1 -> 2 (6)
replace marital_check=1 if marital_combine=="112" | marital_combine=="116"
*marital = 1 -> 2 (6) -> 2 (6)
replace marital_check=1 if marital_combine=="122" | marital_combine=="162" ///
| marital_combine=="126" | marital_combine=="166"
*marital = 1 -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="123" | marital_combine=="163"
*marital = 1 -> 2 (6) -> 4
replace marital_check=1 if marital_combine=="124" | marital_combine=="164"
*marital = 1 -> 2 (6) -> 5
replace marital_check=1 if marital_combine=="125" | marital_combine=="165"
*marital = 2 (6) -> 2 (6) -> 2 (6)
replace marital_check=1 if marital_combine=="222" | marital_combine=="226" ///
| marital_combine=="262" | marital_combine=="622" | ///
marital_combine=="266" | marital_combine=="626" | marital_combine=="662" ///
| marital_combine=="666"
*marital = 2 (6) -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="223" | marital_combine=="663" ///
| marital_combine=="263" | marital_combine=="623"

```

```

*marital = 2 (6) -> 2 (6) -> 4
replace marital_check=1 if marital_combine=="224" | marital_combine=="664" ///
| marital_combine=="264" | marital_combine=="624"
*marital = 2 (6) -> 2 (6) -> 5
replace marital_check=1 if marital_combine=="225" | marital_combine=="665" ///
| marital_combine=="265" | marital_combine=="625"
*marital = 2 (6) -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="232" | marital_combine=="236" ///
| marital_combine=="632" | marital_combine=="636"
*marital = 2 (6) -> 3 -> 3
replace marital_check=1 if marital_combine=="233" | marital_combine=="633"
*marital = 2 (6) -> 4 -> 2 (6)
replace marital_check=1 if marital_combine=="242" | marital_combine=="642" ///
| marital_combine=="246" | marital_combine=="646"
*marital = 2 (6) -> 4 -> 3
replace marital_check=1 if marital_combine=="243" | marital_combine=="643"
*marital = 2 (6) -> 4 -> 4
replace marital_check=1 if marital_combine=="244" | marital_combine=="644"
*marital = 2 (6) -> 5 -> 3
replace marital_check=1 if marital_combine=="253" | marital_combine=="653"
*marital = 2 (6) -> 5 -> 4
replace marital_check=1 if marital_combine=="254" | marital_combine=="654"
*marital = 2 (6) -> 5 -> 5
replace marital_check=1 if marital_combine=="255" | marital_combine=="655"
*marital = 3 -> 3 -> 3
replace marital_check=1 if marital_combine=="333"
*marital = 3 -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="332" | marital_combine=="336"
*marital = 3 -> 2 (6) -> 2 (6)
replace marital_check=1 if marital_combine=="322" | marital_combine=="326" ///
| marital_combine=="362" | marital_combine=="366"
*marital = 3 -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="323" | marital_combine=="363"
*marital = 3 -> 2 (6) -> 4
replace marital_check=1 if marital_combine=="324" | marital_combine=="364"
*marital = 3 -> 2 (6) -> 5
replace marital_check=1 if marital_combine=="325" | marital_combine=="365"
*marital = 4 -> 4 -> 4
replace marital_check=1 if marital_combine=="444"
*marital = 4 -> 4 -> 3
replace marital_check=1 if marital_combine=="443"
*marital = 4 -> 2 (6) -> 2 (6)
replace marital_check=1 if marital_combine=="422" | marital_combine=="426" ///
| marital_combine=="462" | marital_combine=="466"
*marital = 4 -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="423" | marital_combine=="463"
*marital = 4 -> 2 (6) -> 4
replace marital_check=1 if marital_combine=="424" | marital_combine=="464"
*marital = 4 -> 2 (6) -> 5
replace marital_check=1 if marital_combine=="425" | marital_combine=="465"
*marital = 4 -> 3 -> 3
replace marital_check=1 if marital_combine=="433"
*marital = 4 -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="432" | marital_combine=="436"
*marital = 5 -> 5 -> 5
replace marital_check=1 if marital_combine=="555"
*marital = 5 -> 3 -> 3
replace marital_check=1 if marital_combine=="533"
*marital = 5 -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="532" | marital_combine=="536"
*marital = 5 -> 4 -> 4
replace marital_check=1 if marital_combine=="544"
*marital = 5 -> 4 -> 2 (6)
replace marital_check=1 if marital_combine=="542" | marital_combine=="546"

```

```

** number of wave = 4 **
*marital = 1 -> 1 -> 1 -> 1
replace marital_check=1 if marital_combine=="1111"
*marital = 1 -> 1 -> 1 -> 2 (6)
replace marital_check=1 if marital_combine=="1112" | marital_combine=="1116"
*marital = 1 -> 1 -> 2 (6) -> 2 (6)
replace marital_check=1 if marital_combine=="1122" | marital_combine=="1126" ///
| marital_combine=="1162" | marital_combine=="1166"
*marital = 1 -> 1 -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="1123" | marital_combine=="1163"
*marital = 1 -> 1 -> 2 (6) -> 4
replace marital_check=1 if marital_combine=="1124" | marital_combine=="1164"
*marital = 1 -> 1 -> 2 (6) -> 5
replace marital_check=1 if marital_combine=="1125" | marital_combine=="1165"
*marital = 1 -> 2 (6) -> 2 (6) -> 2 (6)
replace marital_check=1 if marital_combine=="1222" | marital_combine=="1226" ///
| marital_combine=="1262" | marital_combine=="1622" | marital_combine=="1266" ///
| marital_combine=="1626" | marital_combine=="1662" | marital_combine=="1666"
*marital = 1 -> 2 (6) -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="1223" | marital_combine=="1263" ///
| marital_combine=="1623" | marital_combine=="1663"
*marital = 1 -> 2 (6) -> 2 (6) -> 4
replace marital_check=1 if marital_combine=="1224" | marital_combine=="1264" ///
| marital_combine=="1624" | marital_combine=="1664"
*marital = 1 -> 2 (6) -> 2 (6) -> 5
replace marital_check=1 if marital_combine=="1225" | marital_combine=="1265" ///
| marital_combine=="1625" | marital_combine=="1665"
*marital = 1 -> 2 (6) -> 3 -> 3
replace marital_check=1 if marital_combine=="1233" | marital_combine=="1633"
*marital = 1 -> 2 (6) -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="1232" | marital_combine=="1236" ///
| marital_combine=="1632" | marital_combine=="1636"
*marital = 1 -> 2 (6) -> 4 -> 4
replace marital_check=1 if marital_combine=="1244" | marital_combine=="1644"
*marital = 1 -> 2 (6) -> 4 -> 2 (6)
replace marital_check=1 if marital_combine=="1242" | marital_combine=="1246" ///
| marital_combine=="1642" | marital_combine=="1646"
*marital = 1 -> 2 (6) -> 4 -> 3
replace marital_check=1 if marital_combine=="1243" | marital_combine=="1643"
*marital = 1 -> 2 (6) -> 5 -> 5
replace marital_check=1 if marital_combine=="1255" | marital_combine=="1643"
*marital = 1 -> 2 (6) -> 5 -> 3
replace marital_check=1 if marital_combine=="1253" | marital_combine=="1263"
*marital = 1 -> 2 (6) -> 5 -> 4
replace marital_check=1 if marital_combine=="1254" | marital_combine=="1654"

*marital = 2 (6) -> 2 (6) -> 2 (6) -> 2 (6)
replace marital_check=1 if marital_combine=="2222" | marital_combine=="2226" ///
| marital_combine=="2262" | marital_combine=="2622" | marital_combine=="6222" ///
| marital_combine=="2266" | marital_combine=="2626" | marital_combine=="6622" ///
| marital_combine=="6262" | marital_combine=="2666" | marital_combine=="6266" ///
| marital_combine=="6662" | marital_combine=="6626" | marital_combine=="6666"
*marital = 2 (6) -> 2 (6) -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="2223" | marital_combine=="6663" ///
| marital_combine=="2663" | marital_combine=="6623" | marital_combine=="6263" ///
| marital_combine=="2263" | marital_combine=="2623" | marital_combine=="6223"
*marital = 2 (6) -> 2 (6) -> 2 (6) -> 4
replace marital_check=1 if marital_combine=="2224" | marital_combine=="6664" ///
| marital_combine=="2664" | marital_combine=="6624" | marital_combine=="6264" ///
| marital_combine=="2264" | marital_combine=="2624" | marital_combine=="6224"
*marital = 2 (6) -> 2 (6) -> 2 (6) -> 5
replace marital_check=1 if marital_combine=="2225" | marital_combine=="6665" ///
| marital_combine=="2665" | marital_combine=="6625" | marital_combine=="6265" ///
| marital_combine=="2265" | marital_combine=="2625" | marital_combine=="6225"
*marital = 2 (6) -> 2 (6) -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="2232" | marital_combine=="2236" ///
| marital_combine=="2632" | marital_combine=="6232" | marital_combine=="2636" ///
| marital_combine=="6236" | marital_combine=="6632" | marital_combine=="6636"
*marital = 2 (6) -> 2 (6) -> 3 -> 3
replace marital_check=1 if marital_combine=="2233" | marital_combine=="6633" ///
| marital_combine=="2633" | marital_combine=="6233"

```

[illegible]

[illegible]


```

*marital = 3 -> 5 -> 5 -> 5
replace marital_check=1 if marital_combine=="3555"
*marital = 4 -> 2 (6) -> 2 (6) -> 2 (6)
replace marital_check=1 if marital_combine=="4222" | marital_combine=="4226" ///
| marital_combine=="4262" | marital_combine=="4622" | marital_combine=="4266" ///
| marital_combine=="4626" | marital_combine=="4662" | marital_combine=="4666"
*marital = 4 -> 2 (6) -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="4223" | marital_combine=="4263" ///
| marital_combine=="4623" | marital_combine=="4663"
*marital = 4 -> 2 (6) -> 2 (6) -> 4
replace marital_check=1 if marital_combine=="4224" | marital_combine=="4264" ///
| marital_combine=="4624" | marital_combine=="4664"
*marital = 4 -> 2 (6) -> 2 (6) -> 5
replace marital_check=1 if marital_combine=="4225" | marital_combine=="4265" ///
| marital_combine=="4625" | marital_combine=="4665"
*marital = 4 -> 2 (6) -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="4232" | marital_combine=="4236" ///
| marital_combine=="4632" | marital_combine=="4636"
*marital = 4 -> 2 (6) -> 3 -> 3
replace marital_check=1 if marital_combine=="4233" | marital_combine=="4633"
*marital = 4 -> 2 (6) -> 4 -> 2 (6)
replace marital_check=1 if marital_combine=="4242" | marital_combine=="4246" ///
| marital_combine=="4642" | marital_combine=="4646"
*marital = 4 -> 2 (6) -> 4 -> 3
replace marital_check=1 if marital_combine=="4243" | marital_combine=="4643"
*marital = 4 -> 2 (6) -> 4 -> 4
replace marital_check=1 if marital_combine=="4244" | marital_combine=="4644"
*marital = 4 -> 2 (6) -> 5 -> 3
replace marital_check=1 if marital_combine=="4253" | marital_combine=="4653"
*marital = 4 -> 2 (6) -> 5 -> 4
replace marital_check=1 if marital_combine=="4254" | marital_combine=="4653"
*marital = 4 -> 2 (6) -> 5 -> 5
replace marital_check=1 if marital_combine=="4255" | marital_combine=="4653"
*marital = 4 -> 3 -> 3 -> 3
replace marital_check=1 if marital_combine=="4333"
*marital = 4 -> 3 -> 2 (6) -> 2 (6)
replace marital_check=1 if marital_combine=="4322" | marital_combine=="4326" ///
| marital_combine=="4362" | marital_combine=="4366"
*marital = 4 -> 3 -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="4323" | marital_combine=="4363"
*marital = 4 -> 3 -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="4332" | marital_combine=="4336"
*marital = 4 -> 3 -> 3 -> 3
replace marital_check=1 if marital_combine=="4333"

```

[illegible]

```

** replace marital_check **
replace marital_check=1 if dum_marital==0
replace marital_check=1 if marital_1=="."
replace marital_check=. if dum_marital==.
drop marital_1 marital_2 marital_3 marital_4 marital_combine
lab var marital_check "marital checking for the same person step3"

** 7.4 - gen dummy for the same individuals with marital status condition **
bysort caseid: gen same_member3=1 if marital_check==1
replace same_member3=0 if same_member3==.
replace same_member3=. if marital_check==.
replace same_member3=. if same_member2==0
lab def same_member3 1 "same" 0 "not same"
lab value same_member3 same_member3
lab var same_member3 "the same individual dummy step3 (marital status condition)"

** 7.5 - check number of unique individuals **

*count unique mameber step 1
by caseid,sort : gen count_memnum1=1 if _n==1 & same_member1==1
sort casehid caseid wave
lab var count_memnum1 "count unique of individual live in more than one wave step1"

*count unique mameber step 2
by caseid,sort : gen count_memnum2=1 if _n==1 & same_member2==1
sort casehid caseid wave
lab var count_memnum2 "count unique of individual live in more than one wave step2"

order same_member1 same_member2 same_member3,after(marital_check)
*-----*

```

- สร้างเงื่อนไขเพื่อตรวจสอบข้อมูลตัวอย่างซ้ำรายบุคคล โดยใช้ caseid เมื่อกำหนดให้คนเดียวกันต้องมีเพศเดียวกัน และมีอายุห่างกันไม่เกิน 1 ปี นับจากการถูกสัมภาษณ์ในครั้งแรกจนถึงครั้งสุดท้าย นอกจากนี้ยังเพิ่มเงื่อนไขของจำนวนปีการศึกษาจะต้องเพิ่มขึ้นไม่เกิน 1 ปี นับจากการถูกสัมภาษณ์ในครั้งแรกจนถึงครั้งสุดท้าย และตั้งชื่อตัวแปรที่สอดคล้องกับเงื่อนไขนี้ว่า same_member4

```

***** STEP 5 check education level *****

** (8) check same member **

** 8.1 - generate education change dummy **
sort casehid caseid wave
/*browse casehid caseid wave round year sex age marital edulevel re_ed same_member1 ///
same_member2 same_member3 same_member4 number_waveid*/
bysort caseid re_ed : egen dum_edu0=sum(count) if same_member2==1
gen dum_edu=number_waveid-dum_edu0
lab var number_waveid "count of wave for panel individuals"
sort casehid caseid wave
replace dum_edu=1 if dum_edu>0 & dum_edu!=.
lab def dum_edu 0 "not change" 1 "change"
lab value dum_edu dum_edu
lab var dum_edu "education level dummy"
drop dum_edu0

** 8.2 - generate year of schooling **
gen ysch=.

*code ysch 2001-2006 */
if year<2007 {
replace ysch = 0 if (edulevel == 0 | edulevel == 1 | edulevel == 2 | edulevel == 3)
replace ysch = 1 if (edulevel == 4)
replace ysch = 2 if (edulevel == 5 | edulevel == 84)
replace ysch = 3 if (edulevel == 6 | edulevel == 85)
replace ysch = 4 if (edulevel == 7 | edulevel == 86)
replace ysch = 5 if (edulevel == 8 | edulevel == 9)
replace ysch = 6 if (edulevel == 10 | edulevel == 35 | edulevel == 87)
replace ysch = 7 if (edulevel == 12 | edulevel == 37)
replace ysch = 8 if (edulevel == 13 | edulevel == 14 | edulevel == 38)
replace ysch = 9 if (edulevel == 15 | edulevel == 39 | edulevel== 88)

```

```

replace ysch = 10 if (edulevel == 17 | edulevel == 40 | edulevel == 89)
replace ysch = 11 if (edulevel == 18 | edulevel == 41)
replace ysch = 12 if (edulevel == 19 | edulevel == 42 | edulevel == 44 ///
| edulevel == 36 | edulevel == 58 | edulevel == 90)
replace ysch = 13 if (edulevel == 11 | edulevel == 21 | edulevel == 23 ///
| edulevel == 29 | edulevel == 78 | edulevel == 91)
replace ysch = 13 if (edulevel == 43 | edulevel == 46 | edulevel == 47)
replace ysch = 13 if (edulevel == 51 | edulevel == 60 | edulevel == 62 ///
| edulevel == 63)
replace ysch = 14 if (edulevel == 16 | edulevel == 22 | edulevel == 24 ///
| edulevel == 48 | edulevel == 45 | edulevel == 59 | edulevel == 79 | edulevel == 92)
replace ysch = 14 if (edulevel == 49 | edulevel == 50 | edulevel == 52 ///
| edulevel == 61)
replace ysch = 14 if (edulevel == 64 | edulevel == 66)
replace ysch = 15 if (edulevel == 20 | edulevel == 25 | edulevel == 53)
replace ysch = 15 if (edulevel == 65 | edulevel == 67 | edulevel == 77 ///
| edulevel == 82 | edulevel == 80 | edulevel == 93)
replace ysch = 16 if (edulevel == 26 | edulevel == 30 | edulevel == 54 ///
| edulevel == 55 | edulevel == 81 | edulevel == 94)
replace ysch = 16 if (edulevel == 56 | edulevel == 57 | edulevel == 68)
replace ysch = 17 if (edulevel == 31)
replace ysch = 17 if (edulevel == 27)
replace ysch = 17 if (edulevel == 30) & (occupptn=="2141" | occupptn=="2224")
replace ysch = 18 if (edulevel == 32 | edulevel == 69)
replace ysch = 18 if (edulevel == 30) & (occupptn=="2221" | occupptn=="2222" | occupptn=="2223")
replace ysch = 21 if (edulevel == 33 | edulevel == 70)
}
else if year>2006 {
/*Change to std. year of schooling 2007-2013*/
replace ysch = 0 if (GRADE_B == 0 | GRADE_B == 110 | GRADE_B == 111 ///
GRADE_B == 001)
replace ysch = 1 if GRADE_B == 211
replace ysch = 2 if GRADE_B == 212
replace ysch = 3 if GRADE_B == 213
replace ysch = 4 if (GRADE_B == 214 | GRADE_B == 219)
replace ysch = 5 if GRADE_B == 215
replace ysch = 6 if (GRADE_B == 216 | GRADE_B == 210 | GRADE_B == 240 ///
| GRADE_B == 250)
replace ysch = 7 if (GRADE_B == 311 | GRADE_B == 321)
replace ysch = 8 if (GRADE_B == 312 | GRADE_B == 322 | GRADE_B == 319 ///
| GRADE_B == 329)
replace ysch = 9 if (GRADE_B == 313 | GRADE_B == 323 | GRADE_B == 310 ///
| GRADE_B == 320 | GRADE_B == 330 | GRADE_B == 340 | GRADE_B == 350)
replace ysch = 10 if (GRADE_B == 411 | GRADE_B == 421)

replace ysch = 11 if (GRADE_B == 412 | GRADE_B == 422 | GRADE_B == 419 ///
| GRADE_B == 429)
replace ysch = 12 if (GRADE_B == 413 | GRADE_B == 423 | GRADE_B == 420 ///
| GRADE_B == 410 | GRADE_B == 430 | GRADE_B == 440 | GRADE_B == 450)
replace ysch = 13 if (GRADE_B == 461 | GRADE_B == 511 | GRADE_B == 519 ///
| GRADE_B == 521 | GRADE_B == 611)
replace ysch = 14 if (GRADE_B == 462 | GRADE_B == 460 | GRADE_B == 510 ///
| GRADE_B == 512 | GRADE_B == 522 | GRADE_B == 529 | GRADE_B == 612 | GRADE_B == 619)
replace ysch = 15 if (GRADE_B == 520 | GRADE_B == 423 | GRADE_B == 613)
replace ysch = 16 if (GRADE_B == 610 | GRADE_B == 614 | GRADE_B == 630 ///
| GRADE_B == 640 | GRADE_B == 650)
replace ysch = 17 if (GRADE_B == 615 | GRADE_B == 660 | GRADE_B == 661 ///
| GRADE_B == 711 | GRADE_B == 719)
replace ysch = 18 if (GRADE_B == 616 | GRADE_B == 710 | GRADE_B == 712)
replace ysch = 19 if (GRADE_B == 760 | GRADE_B == 761 | GRADE_B == 811 ///
| GRADE_B == 871)
replace ysch = 20 if (GRADE_B == 812 | GRADE_B == 819 | GRADE_B == 872)
replace ysch = 21 if (GRADE_B == 810 | GRADE_B == 873 | GRADE_B == 879)
replace ysch = 22 if (GRADE_B == 860 | GRADE_B == 861 | GRADE_B == 874)
replace ysch = 23 if (GRADE_B == 870 | GRADE_B == 875)
}

lab var ysch "Year of Schooling"

```

```

** 8.3 - generate dummy for increasing of year of schooling **
gen ysch_new=ysch
replace ysch_new=. if same_member2==0

bysort caseid: gen ysch0=ysch_new if _n==1 & same_member2==1
bysort caseid: egen ysch1=max(ysch0) if same_member2==1
drop ysch0
gen diff_ysch=ysch_new-ysch1
gen edu_increase0=1 if diff_ysch>=0 & diff_ysch<=1
replace edu_increase0=0 if diff_ysch<0 | diff_ysch>1
replace edu_increase0=1 if diff_ysch==. & same_member2==1
replace edu_increase0=. if same_member2==0
bysort caseid : egen edu_increase=min(edu_increase0) if same_member2==1
lab var edu_increase "checking of year of schooling does not decrease (0-1 year)"
drop ysch1 diff_ysch edu_increase0 ysch_new

** 8.4 - gen dummy for the same individuals with marital status not change **
bysort caseid: gen same_member4=1 if edu_increase==1
replace same_member4=0 if same_member4==.
replace same_member4=. if same_member2==0
lab def same_member4 1 "same" 0 "not same"
lab value same_member4 same_member4
lab var same_member4 "the same individual dummy step4 (check year of schooling)"

```

- สร้างเงื่อนไขเพื่อตรวจสอบข้อมูลตัวอย่างซ้ำรายบุคคล โดยใช้ caseid เมื่อกำหนดให้คนเดียวกันต้องมีเพศเดียวกัน และมีอายุห่างกันไม่เกิน 1 ปี นับจากการถูกสัมภาษณ์ในครั้งแรกจนถึงครั้งสุดท้าย นอกจากนี้ยังเพิ่มเงื่อนไขสถานภาพสมรส และจำนวนปีการศึกษาจะต้องเพิ่มขึ้นไม่เกิน 1 ปี นับจากการถูกสัมภาษณ์ในครั้งแรกจนถึงครั้งสุดท้าย และตั้งชื่อตัวแปรที่สอดคล้องกับเงื่อนไขนี้ว่า same_member5

```

** 8.5 - gen same individuals with marital and education condition
bysort caseid: gen same_member5=1 if same_member3==1 & same_member4==1
replace same_member5=0 if same_member5==.
replace same_member5=. if same_member2==0
lab def same_member5 1 "same" 0 "not same"
lab value same_member5 same_member5
lab var same_member5 "the same individual dummy step5 (check marital status and year of schooling)"

** same_member for IDmatch
gen same_member0 = same_member1
lab def same_member0 1 "same" 0 "not same"
lab value same_member0 same_member0
lab var same_member0 "the same individual dummy (age 0-2) base on caseid"

** order and sort
order wavehid wavehid2 waveid number_waveid dumwaveid count countw_samemem ///
count_memnum1 count_memnum2 total_hhnum total_memnum samenumb dum_enum ///
firstenum lastenum dum_enum numb_members max_numb min_numb group ///
diffage modesex labour head increase marital_check dum_marital dum_edu ysch edu_increase ///
same_member0 same_member1 same_member2 same_member3 same_member4 same_member5,after(enumdist)
sort casehid caseid wave
drop same_member

** fill missing value of same_member with not the same person
replace same_member3=0 if same_member3==.
replace same_member4=0 if same_member4==.
replace same_member5=0 if same_member5==.

```

```
*#####
```

```
**** ..... ****  
**** ..... ****  
**** ..... ****  
**** ..... ****  
**** ..... ****  
**** ..... ****  
**** ..... ID NON MATCHED DATA..... ****  
**** ..... ****  
**** ..... ****  
**** ..... ****  
**** ..... ****  
**** ..... ****  
**** ..... ****  
**** ..... ****  
**** ..... ****  
**** ..... ****  
**** ..... ****  
**** ..... ****  
**** ..... ****  
**** ..... ****  
  
*#####
```

```
***** (4) STEP 7 check same member if age>=0 & age<=2 for ORG NON-MATCHED *****  
  
set more off  
forvalue m=1/4 {  
  
use "${keep}samsetA_rotation`m'_keep.dta",replace  
  
** keep org non-matched data **  
keep if same_member_last==0  
  
** same_member for IDmatch  
gen same_member0 = same_member_last  
lab def same_member0 1 "same" 0 "not same"  
lab value same_member0 same_member0  
lab var same_member0 "the same individual dummy (age 0-2) base on caseid"
```

- สร้างตัวแปรดัมมี่เพื่อตรวจสอบว่า ข้อมูลที่ไม่สามารถเชื่อมโยงได้ด้วย caseid มีสมาชิกครัวเรือนใดที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง เราจะเก็บข้อมูลส่วนดังกล่าวเพื่อนำมาค้นหาและสร้างกลุ่มตัวอย่างซ้ำ และสร้าง caseid ใหม่

```

** keep only household live in more than 1 wave **
bysort casehid wave : gen count_numwave0=1 if _n==1
replace count_numwave0=0 if count_numwave0==.
bysort casehid : egen count_numwave=sum(count_numwave0)
gen countwave_dum=1 if count_numwave>1
replace countwave_dum=0 if count_numwave==1
lab var countwave_dum "checking household live in more than one wave"
drop count_numwave0 count_numwave

** keep household live in more than one wave **
keep if countwave_dum==1

```

- สร้าง caseid ใหม่ โดยเรียงจากเพศและอายุและใช้ wave แรกในการเชื่อมโยงคนเดียวกันใน wave ถัดไป หากสมาชิกคนดังกล่าวมีเพศเดียวกันและมีอายุเท่าเดิมหรือเพิ่มขึ้นไม่เกิน 2 ปีใน wave ถัดไป เราจะสร้าง caseid ใหม่ให้คนดังกล่าวมีหมายเลขเดียวกัน และใช้ชื่อตัวแปรว่า caseid_new และสร้างตัวแปรที่ใช้ตรวจสอบว่าเป็นคนเดียวกันหรือไม่โดยใช้เพศและอายุที่ห่างกันไม่เกิน 2 ปีเป็นเกณฑ์ โดยตั้งชื่อว่า same_member6

```

** (1) re-order of new member id base on first wave **
**gen first wave

gsort casehid wave -age sex

** gen level of rows in first wave **
by casehid : egen firstwave=min(wave)
bysort casehid : gen id=_n if wave==firstwave

** find nerest age and same sex
gen tag0=.
levelsof id, local(levels)
foreach i of local levels {

bysort casehid wave : gen age`i`=age if id==`i' & wave==firstwave
bysort casehid wave : gen sex`i`=sex if id==`i' & wave==firstwave
bysort casehid : egen age`i`_min=min(age`i`)
bysort casehid : egen sex`i`_min=min(sex`i`)
gen tag`i`=`i' if (age-age`i`_min>=0 & age-age`i`_min<=2) & sex`i`_min==sex
replace tag`i`= . if id!=tag`i' & wave==firstwave

bysort casehid wave : gen tag`i`_no=1 if tag`i`!=. & tag=`i`-1'==.
bysort casehid wave : replace tag`i`_no=sum(tag`i`_no) if tag`i`!=.
bysort casehid wave : replace tag`i`= . if tag`i`_no!=1 & tag=`i`-1'==.
bysort casehid wave : replace tag`i`= . if tag=`i`-1'!=.

drop age`i` sex`i` age`i`_min sex`i`_min tag`i`_no
}

levelsof id, local(levels)
foreach i of local levels {
replace tag`i`=0 if tag`i`==.
}

** combine tag and generate newid **
egen newid=rowtotal(tag*)
sort casehid wave sex age
drop tag* id
replace newid=. if newid==0
replace newid=newid+max_numb
tostring newid,gen(newid_s) format(%02.0f)
duplicates tag casehid wave newid_s if newid_s!="",gen(dup)
bysort casehid wave newid_s : replace dup=. if _n!=1 & (dup>0 | dup!=.)
replace newid_s="." if newid_s="" & dup==.
bysort casehid: egen maxid=max(newid)

```

```

** gen new caseid **
gen caseid_new=casehid+newid_s if newid_s!="."
replace caseid_new=caseid if age==.
order casehid caseid caseid_new
lab var caseid_new "NEW CASE IDENTIFICATION NO. SETP 7 NON-MATCHED"
drop firstwave newid newid_s dup

**count wave for caseid_new
replace count=1
bysort caseid_new : egen countwave_idnew=sum(count) if caseid_new!=""
replace countwave_idnew=1 if countwave_idnew==.
sort casehid caseid_new wave
lab var countwave_idnew "checking number of wave by caseid_new"

**gen same individuals with nearest age and same sex
gen same_member6=1 if caseid_new!="" & (countwave_idnew>1)
replace same_member6=0 if same_member6==.
replace same_member6=0 if age==.
lab def same_member6 1 "same" 0 "not same"
lab value same_member6 same_member6
lab var same_member6 "check same member with nearest age and same sex base on first wave (age 0-2)"
*-----*

```

- สร้าง caseid_new โดยเรียงจากเพศและอายุและใช้ wave ที่ 2 ในการเชื่อมโยงคนเดียวกันใน wave ถัดไป และสร้างตัวแปรที่ใช้ตรวจสอบว่าเป็นคนเดียวกันหรือไม่โดยใช้เพศและอายุที่ห่างเกินไม่เกิน 2 ปีเป็นเกณฑ์ โดยตั้งชื่อว่า same_member7

```

** (2) re-order of new member id base on second wave **
** gen level of rows in second wave **
gsort casehid caseid_new wave -age sex
by casehid : egen firstwave =min(wave) if caseid_new==""
bysort casehid : gen id=_n if wave==firstwave & caseid_new==""

** find nerest age and same sex
gen tag0=.
levelsof id, local(levels)
foreach i of local levels {
bysort casehid wave : gen age`i`=age if id==`i' & wave==firstwave
bysort casehid wave : gen sex`i`=sex if id==`i' & wave==firstwave
bysort casehid : egen age`i`_min=min(age`i`)
bysort casehid : egen sex`i`_min=min(sex`i`)

gen tag`i`=`i' if (age-age`i`_min>0 & age-age`i`_min<=2) & sex`i`_min==sex
replace tag`i`= . if id!=tag`i' & wave==firstwave
replace tag`i`= . if caseid_new==""

bysort casehid wave : gen tag`i`_no=1 if tag`i`!=. & tag`i`-1==.
bysort casehid wave : replace tag`i`_no=sum(tag`i`_no) if tag`i`!=.
bysort casehid wave : replace tag`i`= . if tag`i`_no!=1 & tag`i`-1==.
bysort casehid wave : replace tag`i`= . if tag0!=.
drop age`i` sex`i` age`i`_min sex`i`_min tag`i`_no
}

levelsof id, local(levels)
foreach i of local levels {
replace tag`i`=0 if tag`i`==.
}

** combine tag and generate newid **
egen newid=rowtotal(tag*)
sort casehid caseid_new wave sex age
drop tag* id
replace newid=. if newid==0
*add*
replace maxid=0 if maxid==. & caseid_new==""
*--*

```

```

replace newid=newid+maxid
tostring newid,gen(newid_s) format(%02.0f)
capture duplicates tag casehid wave newid_s if newid_s!=".",gen(dup)
capture bysort casehid wave newid_s : replace dup=. if _n!=1 & (dup>0 | dup!=.)
capture replace newid_s="." if newid_s!="" & dup==.
drop maxid
bysort casehid: egen maxid=max(newid)

** gen new caseid **
replace caseid_new=casehid+newid_s if newid_s!="." & caseid_new==""
order casehid caseid caseid_new
lab var caseid_new "NEW CASE INDENTIFICATION NO. SETP 7 NON-MATCHED (diff age 0-2)"
drop firstwave newid newid_s
capture drop dup

**count wave for caseid_new
replace count=1
drop countwave_idnew
bysort caseid_new : egen countwave_idnew=sum(count) if caseid_new!=""
replace countwave_idnew=1 if countwave_idnew==.
sort casehid caseid_new wave
lab var countwave_idnew "checking number of wave by caseid_new"

**gen same individuals with nearest age and same sex
gen same_member7=1 if caseid_new!="." & (countwave_idnew>1)
replace same_member7=0 if same_member7==.
replace same_member7=0 if age==.
lab def same_member7 1 "same" 0 "not same"
lab value same_member7 same_member7
lab var same_member7 "check same member with nearest age and same sex base on second wave (age 0-2)"

*-----*

```

- สร้าง caseid_new โดยเรียงจากเพศและอายุและใช้ wave ที่ 3 ในการเชื่อมโยงคนเดียวกันใน wave ถัดไป และสร้างตัวแปรที่ใช้ตรวจสอบว่าเป็นคนเดียวกันหรือไม่โดยใช้เพศและอายุที่ห่างเกินไม่เกิน 2 ปีเป็นเกณฑ์ โดยตั้งชื่อว่า same_member8

```

gsort casehid caseid_new wave -age sex
by casehid : egen firstwave =min(wave) if caseid_new==""
bysort casehid : gen id=_n if wave==firstwave & caseid_new==""

** find nerest age and same sex
gen tag0=.
levelsof id, local(levels)
foreach i of local(levels) {
  bysort casehid wave : gen age`i`=age if id==`i' & wave==firstwave
  bysort casehid wave : gen sex`i`=sex if id==`i' & wave==firstwave
  bysort casehid : egen age`i`_min=min(age`i`)
  bysort casehid : egen sex`i`_min=min(sex`i`)

  gen tag`i`=`i' if (age-age`i`_min>0 & age-age`i`_min<=2) & sex`i`_min==sex
  replace tag`i`= . if id!=tag`i' & wave==firstwave
  replace tag`i`= . if caseid_new!="."

  bysort casehid wave : gen tag`i`_no=1 if tag`i`!=. & tag`i`-1==.
  bysort casehid wave : replace tag`i`_no=sum(tag`i`_no) if tag`i`!=.
  bysort casehid wave : replace tag`i`= . if tag`i`_no!=1 & tag`i`-1==.
  bysort casehid wave : replace tag`i`= . if tag0!=.
  drop age`i` sex`i` age`i`_min sex`i`_min tag`i`_no
}

levelsof id, local(levels)
foreach i of local(levels) {
  replace tag`i`=0 if tag`i`==.
}

```

```

** combine tag and generate newid **
egen newid=rowtotal(tag*)
sort casehid caseid_new wave sex age
drop tag* id
replace newid=. if newid==0
*add*
replace maxid=0 if maxid==. & caseid_new==""
*---*

replace newid=newid+maxid
tostring newid,gen(newid_s) format(%02.0f)
capture duplicates tag casehid wave newid_s if newid_s!="",gen(dup)
capture bysort casehid wave newid_s : replace dup=. if _n!=1 & (dup>0 | dup!=.)
capture replace newid_s="." if newid_s!="" & dup==.
drop maxid
bysort casehid: egen maxid=max(newid)

** gen new caseid **
replace caseid_new=casehid+newid_s if newid_s!="." & caseid_new==""
order casehid caseid caseid_new
lab var caseid_new "NEW CASE IDENTIFICATION NO. SETP 7 NON-MATCHED (diff age 0-2)"
drop firstwave newid_s
capture drop dup

**count wave for caseid_new
replace count=1
drop countwave_idnew
bysort caseid_new : egen countwave_idnew=sum(count) if caseid_new!=""
replace countwave_idnew=1 if countwave_idnew==.
sort casehid caseid_new wave
lab var countwave_idnew "checking number of wave by caseid_new"

**gen same individuals with nearest age and same sex
gen same_member8=1 if caseid_new!="" & (countwave_idnew>1)
replace same_member8=0 if same_member8==.
replace same_member8=0 if age==.
lab def same_member8 1 "same" 0 "not same"
lab value same_member8 same_member8
lab var same_member8 "check same member with nearest age and same sex base on third wave (age 0-2)"

*-----*

```

- สร้าง caseid_new โดยเรียงจากเพศและอายุและใช้ wave ที่ 4 ในการเชื่อมโยงคนเดียวกันใน wave ถัดไป และสร้างตัวแปรที่ใช้ตรวจสอบว่าเป็นคนเดียวกันหรือไม่โดยใช้เพศและอายุที่ห่างเกินไม่เกิน 2 ปีเป็นเกณฑ์ โดยตั้งชื่อว่า same_member9

```

** (4) re-order of new member id base on forth wave **
** gen level of rows in forth wave **
gsort casehid caseid_new wave -age sex
by casehid : egen firstwave =min(wave) if caseid_new==""
bysort casehid : gen id=_n if wave==firstwave & caseid_new==""

** find nerest age and same sex
gen tag0=.
levelsof id, local(levels)
foreach i of local(levels) {
  bysort casehid wave : gen age`i`=age if id==`i' & wave==firstwave
  bysort casehid wave : gen sex`i`=sex if id==`i' & wave==firstwave
  bysort casehid : egen age`i`_min=min(age`i`)
  bysort casehid : egen sex`i`_min=min(sex`i`)

  gen tag`i`=`i' if (age-age`i`_min>=0 & age-age`i`_min<=2) & sex`i`_min==sex
  replace tag`i`= . if id!=tag`i' & wave==firstwave
  replace tag`i`= . if caseid_new!=""
}

```

```

bysort casehid wave : gen tag`i'_no=1 if tag`i'!=. & tag`i'-1'==.
bysort casehid wave : replace tag`i'_no=sum(tag`i'_no) if tag`i'!=.
bysort casehid wave : replace tag`i'=. if tag`i'_no!=1 & tag`i'-1'==.
bysort casehid wave : replace tag`i'=. if tag0!=.
drop age`i' sex`i' age`i'_min sex`i'_min tag`i'_no
}

levelsof id, local(levels)
foreach i of local levels {
  replace tag`i'=0 if tag`i'==.
}

** combine tag and generate newid **
egen newid=rowtotal(tag*)
sort casehid caseid_new wave sex age
drop tag* id
replace newid=. if newid==0
*add*
replace maxid=0 if maxid==. & caseid_new=="
*---*

replace newid=newid+maxid
tostring newid,gen(newid_s) format(%02.0f)
capture duplicates tag casehid wave newid_s if newid_s!="",gen(dup)
capture bysort casehid wave newid_s : replace dup=. if _n!=1 & (dup>0 | dup!=.)
capture replace newid_s="." if newid_s!="" & dup==.
drop maxid
bysort casehid: egen maxid=max(newid)

** gen new caseid **
replace caseid_new=casehid+newid_s if newid_s!="." & caseid_new=="
order casehid caseid caseid_new
lab var caseid_new "NEW CASE IDENTIFICATION NO. STEP 7 NON-MATCHED (diff age 0-2)"
drop firstwave newid newid_s
capture drop dup

**count wave for caseid_new
replace count=1
drop countwave_idnew
bysort caseid_new : egen countwave_idnew=sum(count) if caseid_new!=""
replace countwave_idnew=1 if countwave_idnew==.
sort casehid caseid new wave
lab var countwave_idnew "checking number of wave by caseid_new"

**gen same individuals with nearest age and same sex
gen same_member9=1 if caseid_new!="" & (countwave_idnew>1)
replace same_member9=0 if same_member9==.
replace same_member9=0 if age==.
lab def same_member9 1 "same" 0 "not same"
lab value same_member9 same_member9
lab var same_member9 "check same member with nearest age and same sex base on forth wave (age 0-2)"

**combine same_member in nonmatched data with matched data
gen same_member1=same_member9
*-----*

** re-order variables
drop same_member_last
lab var same_member1 "the same individual dummy step1 (age 0-2)"
lab var maxid "maximun member number in step7"
drop diffage modesex group increase dumwaveid
order maxid count* countw_samemem countwave_dum countwave_idnew same_member ///
same_member1 same_member6 same_member7 same_member8 same_member9,after( samenumb )

*count unique mameber step 7
by caseid_new,sort : gen count_memnum3=1 if _n==1
sort casehid caseid wave
lab var count_memnum3 "count unique of individual live in more than one wave step7"
*****

```

- สร้าง caseid ใหม่ โดยเรียงจากเพศและอายุและใช้ wave แรกในการเชื่อมโยงคนเดียวกันใน wave ถัดไป หากสมาชิกคนดังกล่าวมีเพศเดียวกันและมีอายุเท่าเดิมหรือเพิ่มขึ้นไม่เกิน 1 ปีใน wave ถัดไป เราจะสร้าง caseid ใหม่ให้คนดังกล่าวมีหมายเลขเดียวกัน และใช้ชื่อตัวแปรว่า caseid_new2 และสร้างตัวแปรที่ใช้ตรวจสอบว่าเป็นคนเดียวกันหรือไม่โดยใช้เพศและอายุที่ห่างเกินไม่เกิน 1 ปีเป็นเกณฑ์ โดยตั้งชื่อว่า same_member10

```
***** STEP 8 check same member if age>=0 & age<=1 for ORG NON-MATCHED DATA *****

** (1) re-order of new member id base on first wave **
**gen first wave
drop maxid
gsort casehid wave -age sex

** gen level of rows in first wave **
by casehid : egen firstwave=min(wave)
bysort casehid : gen id=_n if wave==firstwave

** find nerest age and same sex
gen tag0=.
levelsof id, local(levels)
foreach i of local levels {

bysort casehid wave : gen age`i`=age if id==`i' & wave==firstwave
bysort casehid wave : gen sex`i`=sex if id==`i' & wave==firstwave
bysort casehid : egen age`i`_min=min(age`i`)
bysort casehid : egen sex`i`_min=min(sex`i`)
gen tag`i`=`i' if (age-age`i`_min>=0 & age-age`i`_min<=1) & sex`i`_min==sex
replace tag`i`= . if id!=tag`i' & wave==firstwave

bysort casehid wave : gen tag`i`_no=1 if tag`i`!=. & tag=`i`-1'==.
bysort casehid wave : replace tag`i`_no=sum(tag`i`_no) if tag`i`!=.
bysort casehid wave : replace tag`i`= . if tag`i`_no!=1 & tag=`i`-1'==.
bysort casehid wave : replace tag`i`= . if tag=`i`-1'!=.

drop age`i` sex`i` age`i`_min sex`i`_min tag`i`_no
}

levelsof id, local(levels)
foreach i of local levels {
replace tag`i`=0 if tag`i`==.
}

** combine tag and generate newid **
egen newid=rowtotal(tag*)
sort casehid wave sex age
drop tag* id
replace newid=. if newid==0
replace newid=newid+max_numb
tostring newid,gen(newid_s) format(%02.0f)
duplicates tag casehid wave newid_s if newid_s!="",gen(dup)
bysort casehid wave newid_s : replace dup=. if _n!=1 & (dup>0 | dup!=.)
replace newid_s="." if newid_s="" & dup==.
bysort casehid: egen maxid=max(newid)

** gen new caseid2 (diff age 0-1) **
gen caseid_new2=casehid+newid_s if newid_s!="."
replace caseid_new2=caseid if age==.
order casehid caseid caseid_new caseid_new2
lab var caseid_new2 "NEW CASE INDENTIFICATION NO. SETP 8 NON-MATCHED (diff age 0-1)"
drop firstwave newid newid_s dup
```

```

**count wave for caseid_new2
replace count=1
bysort caseid_new2 : egen countwave_idnew2=sum(count) if caseid_new2!="
replace countwave_idnew2=1 if countwave_idnew2==.
sort casehid caseid_new2 wave
lab var countwave_idnew2 "checking number of wave by caseid_new2"

**gen same individuals with nearest age and same sex
gen same_member10=1 if caseid_new2!=" & (countwave_idnew2>1)
replace same_member10=0 if same_member10==.
replace same_member10=0 if age==.
lab def same_member10 1 "same" 0 "not same"
lab value same_member10 same_member10
lab var same_member10 "check same member with nearest age and same sex base on first wave (age 0-1)"
*-----*

```

- สร้าง caseid_new2 โดยเรียงจากเพศและอายุและใช้ wave ที่ 2 ในการเชื่อมโยงคนเดียวกันใน wave ถัดไป และสร้างตัวแปรที่ใช้ตรวจสอบว่าเป็นคนเดียวกันหรือไม่โดยใช้เพศและอายุที่ห่างเกินไม่เกิน 1 ปีเป็นเกณฑ์ โดยตั้งชื่อว่า same_member11

```

** (2) re-order of new member id base on second wave **
** gen level of rows in second wave **
gsort casehid caseid_new2 wave -age sex
by casehid : egen firstwave =min(wave) if caseid_new2=="
bysort casehid : gen id=_n if wave==firstwave & caseid_new2=="

** find nerest age and same sex
gen tag0=.
levelsof id, local(levels)
foreach i of local levels {
bysort casehid wave : gen age`i`=age if id==`i' & wave==firstwave
bysort casehid wave : gen sex`i`=sex if id==`i' & wave==firstwave
bysort casehid : egen age`i`_min=min(age`i`)
bysort casehid : egen sex`i`_min=min(sex`i`)

gen tag`i`=`i' if (age-age`i`_min>=0 & age-age`i`_min<=1) & sex`i`_min==sex
replace tag`i`= . if id!=tag`i' & wave==firstwave
replace tag`i`= . if caseid_new2!="

bysort casehid wave : gen tag`i`_no=1 if tag`i`!=. & tag=`i`-1'==.
bysort casehid wave : replace tag`i`_no=sum(tag`i`_no) if tag`i`!=.
bysort casehid wave : replace tag`i`= . if tag`i`_no!=1 & tag=`i`-1'==.
bysort casehid wave : replace tag`i`= . if tag0!=.
drop age`i` sex`i` age`i`_min sex`i`_min tag`i`_no
}

levelsof id, local(levels)
foreach i of local levels {
replace tag`i`=0 if tag`i`==.
}

** combine tag and generate newid **
egen newid=rowtotal(tag*)
sort casehid caseid_new2 wave sex age
drop tag* id
replace newid=. if newid==0
*add*
replace maxid=0 if maxid==. & caseid_new2=="
*---*

```

```

replace newid=newid+maxid
tostring newid,gen(newid_s) format(%02.0f)
capture duplicates tag casehid wave newid_s if newid_s!=".",gen(dup)
capture bysort casehid wave newid_s : replace dup=. if _n!=1 & (dup>0 | dup!=.)
capture replace newid_s="." if newid_s!="" & dup==.
drop maxid
bysort casehid: egen maxid=max(newid)

** gen new caseid2 **
replace caseid_new2=casehid+newid_s if newid_s!="." & caseid_new2==""
order casehid caseid caseid_new2
lab var caseid_new2 "NEW CASE IDENTIFICATION NO. SETP 8 NON-MATCHED (diff age 0-1)"
drop firstwave newid newid_s
capture drop dup

**count wave for caseid_new
replace count=1
drop countwave_idnew2
bysort caseid_new2 : egen countwave_idnew2=sum(count) if caseid_new2!=""
replace countwave_idnew2=1 if countwave_idnew2==.
sort casehid caseid_new2 wave
lab var countwave_idnew2 "checking number of wave by caseid_new2"

**gen same individuals with nearest age and same sex
gen same_member11=1 if caseid_new2!="" & (countwave_idnew2>1)
replace same_member11=0 if same_member11==.
replace same_member11=0 if age==.
lab def same_member11 1 "same" 0 "not same"
lab value same_member11 same_member11
lab var same_member11 "check same member with nearest age and same sex base on second wave (age 0-1)"
*-----*

```

- สร้าง caseid_new2 โดยเรียงจากเพศและอายุและใช้ wave ที่ 3 ในการเชื่อมโยงคนเดียวกัน
ใน wave ถัดไป และสร้างตัวแปรที่ใช้ตรวจสอบว่าเป็นคนเดียวกันหรือไม่โดยใช้เพศและอายุที่ห่างเกิน
ไม่ เกิน 1 ปีเป็นเกณฑ์ โดยตั้งชื่อว่า same_member12

```

** (3) re-order of new member id base on third wave **
** gen level of rows in third wave **
gsort casehid caseid_new2 wave -age sex
by casehid : egen firstwave =min(wave) if caseid_new2==""
bysort casehid : gen id=_n if wave==firstwave & caseid_new2==""

** find nerest age and same sex
gen tag0=.
levelsof id, local(levels)
foreach i of local levels {
bysort casehid wave : gen age`i`=age if id==`i' & wave==firstwave
bysort casehid wave : gen sex`i`=sex if id==`i' & wave==firstwave
bysort casehid : egen age`i`_min=min(age`i`)
bysort casehid : egen sex`i`_min=min(sex`i`)

gen tag`i`=`i' if (age-age`i`_min>=0 & age-age`i`_min<=1) & sex`i`_min==sex
replace tag`i`= . if id!=tag`i' & wave==firstwave
replace tag`i`= . if caseid_new2!="

bysort casehid wave : gen tag`i`_no=1 if tag`i`!=. & tag=`i`-1'==.
bysort casehid wave : replace tag`i`_no=sum(tag`i`_no) if tag`i`!=.
bysort casehid wave : replace tag`i`= . if tag`i`_no!=1 & tag=`i`-1'==.
bysort casehid wave : replace tag`i`= . if tag0!=.
drop age`i` sex`i` age`i`_min sex`i`_min tag`i`_no
}

levelsof id, local(levels)
foreach i of local levels {
replace tag`i`=0 if tag`i`==.
}

```

```

** combine tag and generate newid **
egen newid=rowtotal(tag*)
sort casehid caseid_new2 wave sex age
drop tag* id
replace newid=. if newid==0
*add*
replace maxid=0 if maxid==. & caseid_new2=="
*---*

replace newid=newid+maxid
tostring newid,gen(newid_s) format(%02.0f)
capture duplicates tag casehid wave newid_s if newid_s!="",gen(dup)
capture bysort casehid wave newid_s : replace dup=. if _n!=1 & (dup>0 | dup!=.)
capture replace newid_s="." if newid_s!="" & dup==.
drop maxid
bysort casehid: egen maxid=max(newid)

** gen new caseid2 **
replace caseid_new2=casehid+newid_s if newid_s!="." & caseid_new2=="
order casehid caseid caseid_new2
lab var caseid_new2 "NEW CASE IDENTIFICATION NO. SETP 8 NON-MATCHED (diff age 0-1)"
drop firstwave newid newid_s
capture drop dup

**count wave for caseid_new2
replace count=1
drop countwave_idnew2
bysort caseid_new2 : egen countwave_idnew2=sum(count) if caseid_new2!="
replace countwave_idnew2=1 if countwave_idnew2==.
sort casehid caseid_new2 wave
lab var countwave_idnew2 "checking number of wave by caseid_new2"

**gen same individuals with nearest age and same sex
gen same_member12=1 if caseid_new2!="." & (countwave_idnew2>1)
replace same_member12=0 if same_member8==.
replace same_member12=0 if age==.
lab def same_member12 1 "same" 0 "not same"
lab value same_member12 same_member12
lab var same_member12 "check same member with nearest age and same sex base on third wave (age 0-1)"

*-----*

```

- สร้าง caseid_new2 โดยเรียงจากเพศและอายุและใช้ wave ที่ 4 ในการเชื่อมโยงคนเดียวกันใน wave ถัดไป และสร้างตัวแปรที่ใช้ตรวจสอบว่าเป็นคนเดียวกันหรือไม่โดยใช้เพศและอายุที่ห่างเกินไม่เกิน 1 ปีเป็นเกณฑ์ โดยตั้งชื่อว่า same_member13

```

** (4) re-order of new member id base on forth wave **
** gen level of rows in forth wave **
gsort casehid caseid_new2 wave -age sex
by casehid : egen firstwave =min(wave) if caseid_new2=="
bysort casehid : gen id=_n if wave==firstwave & caseid_new2=="

** find nerest age and same sex
gen tag0=.
levelsof id, local(levels)
foreach i of local(levels) {
  bysort casehid wave : gen age`i`=age if id==`i' & wave==firstwave
  bysort casehid wave : gen sex`i`=sex if id==`i' & wave==firstwave
  bysort casehid : egen age`i`_min=min(age`i`)
  bysort casehid : egen sex`i`_min=min(sex`i`)

  gen tag`i`=`i' if (age-age`i`_min>=0 & age-age`i`_min<=1) & sex`i`_min==sex
  replace tag`i`= . if id!=tag`i' & wave==firstwave
  replace tag`i`= . if caseid_new2!="

```

```

bysort casehid wave : gen tag`i'_no=1 if tag`i'!=. & tag`i'-1'==.
bysort casehid wave : replace tag`i'_no=sum(tag`i'_no) if tag`i'!=.
bysort casehid wave : replace tag`i'=1 if tag`i'_no!=1 & tag`i'-1'==.
bysort casehid wave : replace tag`i'=0 if tag`i'_no==0.
drop age`i' sex`i' age`i'_min sex`i'_min tag`i'_no
}

levelsof id, local(levels)
foreach i of local levels {
  replace tag`i'=0 if tag`i'==.
}

** combine tag and generate newid **
egen newid=rowtotal(tag*)
sort casehid caseid_new2 wave sex age
drop tag* id
replace newid=. if newid==0
*add*
replace maxid=0 if maxid==. & caseid_new2=="
*---*

replace newid=newid+maxid
tostring newid,gen(newid_s) format(%02.0f)
capture duplicates tag casehid wave newid_s if newid_s!="",gen(dup)
capture bysort casehid wave newid_s : replace dup=. if _n!=1 & (dup>0 | dup!=.)
capture replace newid_s="." if newid_s==" " & dup==.
drop maxid
bysort casehid: egen maxid=max(newid)

** gen new caseid **
replace caseid_new2=casehid+newid_s if newid_s!=" " & caseid_new2==" "
order casehid caseid caseid_new2
lab var caseid_new2 "NEW CASE IDENTIFICATION NO. SETP 8 NON-MATCHED (diff age 0-1)"
drop firstwave newid newid_s
capture drop dup

**count wave for caseid_new2
replace count=1
drop countwave_idnew2
bysort caseid_new2 : egen countwave_idnew2=sum(count) if caseid_new2!=" "
replace countwave_idnew2=1 if countwave_idnew2==.
sort casehid caseid_new2 wave
lab var countwave_idnew2 "checking number of wave by caseid_new2"

**gen same individuals with nearest age and same sex
gen same_member13=1 if caseid_new2!=" " & (countwave_idnew2>1)
replace same_member13=0 if same_member13==.
replace same_member13=0 if age==.
lab def same_member13 1 "same" 0 "not same"
lab value same_member13 same_member13
lab var same_member13 "check same member with nearest age and same sex base on forth wave (age 0-1)"

**combine same_member in nonmatched data with matched data
gen same_member2=same_member13
*-----*

** re-order variables
lab var maxid "maximun member number in step8"
order maxid count* countw_samemem countwave_dum countwave_idnew same_member*,after( samenumb )

*count unique mameber step 8
by caseid_new2,sort : gen count_memnum4=1 if _n==1
sort casehid caseid_new2 wave
lab var count_memnum4 "count unique of individual live in more than one wave step8"

*****

```

- สร้างเงื่อนไขเพื่อตรวจสอบข้อมูลตัวอย่างซ้ำรายบุคคล โดยใช้ caseid_new2 เมื่อกำหนดให้คนเดียวกันต้องมีเพศเดียวกัน และมีอายุห่างกันไม่เกิน 1 ปีนับจากการถูกสัมภาษณ์ครั้งแรกจนถึงครั้งสุดท้าย นอกจากนี้ยังเพิ่มเงื่อนไขของสถานภาพสมรส และตั้งชื่อตัวแปรที่สอดคล้องกับเงื่อนไขนี้ว่า same_member14

```

***** STEP 9 check marital status and marital status change *****
** 1.1 - generate marital change dummy **
*browse casehid caseid_new wave round year sex age marital same_member1 same_member2
*drop number_waveid
bysort caseid_new2 : egen number_waveid0=sum(count) if same_member13==1
bysort caseid_new2 : egen number_waveid=max(number_waveid0)
drop number_waveid0
sort casehid caseid_new2 wave
bysort caseid_new2 marital: egen dum_marital0=sum(count) if same_member13==1
replace number_waveid=. if dum_marital0==.
gen dum_marital=number_waveid - dum_marital0
replace dum_marital=1 if dum_marital>0 & dum_marital!=.
lab def dum_marital 0 "not change" 1 "change"
lab value dum_marital dum_marital
lab var dum_marital "marital status dummy"
drop dum_marital0
*tab dum_marital if count_memnum==1

** 1.2 - checking marital status by wave **
sort casehid caseid_new2 wave
*wave1
bysort caseid_new2 : gen marital_0=marital if _n==1 & dum_marital!=.
bysort caseid_new2 : egen marital_1=max(marital_0)
drop marital_0
*wave2
bysort caseid_new2 : gen marital_0=marital if _n==2 & dum_marital!=.
bysort caseid_new2 : egen marital_2=max(marital_0)
drop marital_0
*wave3
bysort caseid_new2 : gen marital_0=marital if _n==3 & dum_marital!=.
bysort caseid_new2 : egen marital_3=max(marital_0)
drop marital_0
*wave4
bysort caseid_new2 : gen marital_0=marital if _n==4 & dum_marital!=.
bysort caseid_new2 : egen marital_4=max(marital_0)
drop marital_0

**replace dum_marital=not change if age wave1 < 15 (marital_1==.)
replace dum_marital=0 if marital_1==.

** 1.3 - generate same person with marital status condition **
tostring marital_1 marital_2 marital_3 marital_4,replace

gen str marital_combine=""
replace marital_combine=marital_1+marital_2+marital_3+marital_4
replace marital_combine = subinstr(marital_combine, ".", "",.)

gen marital_check=0

** number of wave = 2 **
*marital = 1 -> 1,2 (6)
replace marital_check=1 if marital_combine=="11" | marital_combine=="12" ///
| marital_combine=="16"
*marital = 2 -> 2(6),3,4,5
replace marital_check=1 if marital_combine=="22" | marital_combine=="26" ///
| marital_combine=="23" | marital_combine=="24" | marital_combine=="25"

```

```

*marital = 3 -> 2(6),3
replace marital_check=1 if marital_combine=="32" | marital_combine=="36" ///
| marital_combine=="33"
*marital = 4 -> 2(6),3,4
replace marital_check=1 if marital_combine=="42" | marital_combine=="46" ///
| marital_combine=="43" | marital_combine=="44"
*marital = 5 -> 3,4,5
replace marital_check=1 if marital_combine=="53" | marital_combine=="54" ///
| marital_combine=="55"

** number of wave = 3 **
*marital = 1 -> 1 -> 1
replace marital_check=1 if marital_combine=="111"
*marital = 1 -> 1 -> 2 (6)
replace marital_check=1 if marital_combine=="112" | marital_combine=="116"
*marital = 1 -> 2 (6) -> 2 (6)
replace marital_check=1 if marital_combine=="122" | marital_combine=="162" ///
| marital_combine=="126" | marital_combine=="166"
*marital = 1 -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="123" | marital_combine=="163"
*marital = 1 -> 2 (6) -> 4
replace marital_check=1 if marital_combine=="124" | marital_combine=="164"
*marital = 1 -> 2 (6) -> 5
replace marital_check=1 if marital_combine=="125" | marital_combine=="165"
*marital = 2 (6) -> 2 (6) -> 2 (6)
replace marital_check=1 if marital_combine=="222" | marital_combine=="226" ///
| marital_combine=="262" | marital_combine=="622" | marital_combine=="266" ///
| marital_combine=="626" | marital_combine=="662" | marital_combine=="666"
*marital = 2 (6) -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="223" | marital_combine=="663" ///
| marital_combine=="263" | marital_combine=="623"
*marital = 2 (6) -> 2 (6) -> 4
replace marital_check=1 if marital_combine=="224" | marital_combine=="664" ///
| marital_combine=="264" | marital_combine=="624"
*marital = 2 (6) -> 2 (6) -> 5
replace marital_check=1 if marital_combine=="225" | marital_combine=="665" ///
| marital_combine=="265" | marital_combine=="625"
*marital = 2 (6) -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="232" | marital_combine=="236" ///
| marital_combine=="632" | marital_combine=="636"
*marital = 2 (6) -> 3 -> 3
replace marital_check=1 if marital_combine=="233" | marital_combine=="633"
*marital = 2 (6) -> 4 -> 3
replace marital_check=1 if marital_combine=="243" | marital_combine=="643"
*marital = 2 (6) -> 4 -> 4
replace marital_check=1 if marital_combine=="244" | marital_combine=="644"
*marital = 2 (6) -> 5 -> 3
replace marital_check=1 if marital_combine=="253" | marital_combine=="653"
*marital = 2 (6) -> 5 -> 4
replace marital_check=1 if marital_combine=="254" | marital_combine=="654"
*marital = 2 (6) -> 5 -> 5
replace marital_check=1 if marital_combine=="255" | marital_combine=="655"
*marital = 3 -> 3 -> 3
replace marital_check=1 if marital_combine=="333"
*marital = 3 -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="332" | marital_combine=="336"
*marital = 3 -> 2 (6) -> 2 (6)
replace marital_check=1 if marital_combine=="322" | marital_combine=="326" ///
| marital_combine=="362" | marital_combine=="366"
*marital = 3 -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="323" | marital_combine=="363"
*marital = 3 -> 2 (6) -> 4
replace marital_check=1 if marital_combine=="324" | marital_combine=="364"
*marital = 3 -> 2 (6) -> 5
replace marital_check=1 if marital_combine=="325" | marital_combine=="365"
*marital = 4 -> 4 -> 4
replace marital_check=1 if marital_combine=="444"

```

```

*marital = 4 -> 4 -> 3
replace marital_check=1 if marital_combine=="443"
*marital = 4 -> 2 (6) -> 2 (6)
replace marital_check=1 if marital_combine=="422" | marital_combine=="426" ///
| marital_combine=="462" | marital_combine=="466"
*marital = 4 -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="423" | marital_combine=="463"
*marital = 4 -> 2 (6) -> 4
replace marital_check=1 if marital_combine=="424" | marital_combine=="464"
*marital = 4 -> 2 (6) -> 5
replace marital_check=1 if marital_combine=="425" | marital_combine=="465"
*marital = 4 -> 3 -> 3
replace marital_check=1 if marital_combine=="433"
*marital = 4 -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="432" | marital_combine=="436"
*marital = 5 -> 5 -> 5
replace marital_check=1 if marital_combine=="555"
*marital = 5 -> 3 -> 3
replace marital_check=1 if marital_combine=="533"
*marital = 5 -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="532" | marital_combine=="536"
*marital = 5 -> 4 -> 4
replace marital_check=1 if marital_combine=="544"
*marital = 5 -> 4 -> 2 (6)
replace marital_check=1 if marital_combine=="542" | marital_combine=="546"

** number of wave = 4 **
*marital = 1 -> 1 -> 1 -> 1
replace marital_check=1 if marital_combine=="1111"
*marital = 1 -> 1 -> 1 -> 2 (6)
replace marital_check=1 if marital_combine=="1112" | marital_combine=="1116"
*marital = 1 -> 1 -> 2 (6) -> 2 (6)
replace marital_check=1 if marital_combine=="1122" | marital_combine=="1126" ///
| marital_combine=="1162" | marital_combine=="1166"
*marital = 1 -> 1 -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="1123" | marital_combine=="1163"
*marital = 1 -> 1 -> 2 (6) -> 4
replace marital_check=1 if marital_combine=="1124" | marital_combine=="1164"
*marital = 1 -> 1 -> 2 (6) -> 5
replace marital_check=1 if marital_combine=="1125" | marital_combine=="1165"
*marital = 1 -> 2 (6) -> 2 (6) -> 2 (6)
replace marital_check=1 if marital_combine=="1222" | marital_combine=="1226" ///
| marital_combine=="1262" | marital_combine=="1622" | marital_combine=="1266" ///
| marital_combine=="1626" | marital_combine=="1662" | marital_combine=="1666"
*marital = 1 -> 2 (6) -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="1223" | marital_combine=="1263" ///
| marital_combine=="1623" | marital_combine=="1663"
*marital = 1 -> 2 (6) -> 2 (6) -> 4
replace marital_check=1 if marital_combine=="1224" | marital_combine=="1264" ///
| marital_combine=="1624" | marital_combine=="1664"
*marital = 1 -> 2 (6) -> 2 (6) -> 5
replace marital_check=1 if marital_combine=="1225" | marital_combine=="1265" ///
| marital_combine=="1625" | marital_combine=="1665"
*marital = 1 -> 2 (6) -> 3 -> 3
replace marital_check=1 if marital_combine=="1233" | marital_combine=="1633"
*marital = 1 -> 2 (6) -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="1232" | marital_combine=="1236" ///
| marital_combine=="1632" | marital_combine=="1636"
*marital = 1 -> 2 (6) -> 4 -> 4
replace marital_check=1 if marital_combine=="1244" | marital_combine=="1644"
*marital = 1 -> 2 (6) -> 4 -> 2 (6)
replace marital_check=1 if marital_combine=="1242" | marital_combine=="1246" ///
| marital_combine=="1642" | marital_combine=="1646"

```

```

*marital = 1 -> 2 (6) -> 4 -> 3
replace marital_check=1 if marital_combine=="1243" | marital_combine=="1643"
*marital = 1 -> 2 (6) -> 5 -> 5
replace marital_check=1 if marital_combine=="1255" | marital_combine=="1643"
*marital = 1 -> 2 (6) -> 5 -> 3
replace marital_check=1 if marital_combine=="1253" | marital_combine=="1263"
*marital = 1 -> 2 (6) -> 5 -> 4
replace marital_check=1 if marital_combine=="1254" | marital_combine=="1654"
*marital = 2 (6) -> 2 (6) -> 2 (6) -> 2 (6)
replace marital_check=1 if marital_combine=="2222" | marital_combine=="2226" ///
| marital_combine=="2262" | marital_combine=="2622" | marital_combine=="6222" ///
| marital_combine=="2266" | marital_combine=="2626" | marital_combine=="6622" ///
| marital_combine=="6262" | marital_combine=="2666" | marital_combine=="6266" ///
| marital_combine=="6662" | marital_combine=="6626" | marital_combine=="6666"
*marital = 2 (6) -> 2 (6) -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="2223" | marital_combine=="6663" ///
| marital_combine=="2663" | marital_combine=="6623" | marital_combine=="6263" ///
| marital_combine=="2263" | marital_combine=="2623" | marital_combine=="6223"
*marital = 2 (6) -> 2 (6) -> 2 (6) -> 4
replace marital_check=1 if marital_combine=="2224" | marital_combine=="6664" ///
| marital_combine=="2664" | marital_combine=="6624" | marital_combine=="6264" ///
| marital_combine=="2264" | marital_combine=="2624" | marital_combine=="6224"
*marital = 2 (6) -> 2 (6) -> 2 (6) -> 5
replace marital_check=1 if marital_combine=="2225" | marital_combine=="6665" ///
| marital_combine=="2665" | marital_combine=="6625" | marital_combine=="6265" ///
| marital_combine=="2265" | marital_combine=="2625" | marital_combine=="6225"
*marital = 2 (6) -> 2 (6) -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="2232" | marital_combine=="2236" ///
| marital_combine=="2632" | marital_combine=="6232" | marital_combine=="2636" ///
| marital_combine=="6236" | marital_combine=="6632" | marital_combine=="6636"
*marital = 2 (6) -> 2 (6) -> 3 -> 3
replace marital_check=1 if marital_combine=="2233" | marital_combine=="6633" ///
| marital_combine=="2633" | marital_combine=="6233"
*marital = 2 (6) -> 2 (6) -> 4 -> 2 (6)
replace marital_check=1 if marital_combine=="2242" | marital_combine=="2246" ///
| marital_combine=="2642" | marital_combine=="6242" | marital_combine=="2646" ///
| marital_combine=="6246" | marital_combine=="6642" | marital_combine=="6646"
*marital = 2 (6) -> 2 (6) -> 4 -> 3
replace marital_check=1 if marital_combine=="2243" | marital_combine=="2643" ///
| marital_combine=="6243" | marital_combine=="6643"
*marital = 2 (6) -> 2 (6) -> 4 -> 4
replace marital_check=1 if marital_combine=="2244" | marital_combine=="2644" ///
| marital_combine=="6244" | marital_combine=="6644"
*marital = 2 (6) -> 2 (6) -> 5 -> 3
replace marital_check=1 if marital_combine=="2253" | marital_combine=="2653" ///
| marital_combine=="6253" | marital_combine=="6653"
*marital = 2 (6) -> 2 (6) -> 5 -> 4
replace marital_check=1 if marital_combine=="2254" | marital_combine=="2654" ///
| marital_combine=="6254" | marital_combine=="6654"
*marital = 2 (6) -> 2 (6) -> 5 -> 5
replace marital_check=1 if marital_combine=="2255" | marital_combine=="2655" ///
| marital_combine=="6255" | marital_combine=="6655"
*marital = 2 (6) -> 3 -> 2 (6) -> 2 (6)
replace marital_check=1 if marital_combine=="2322" | marital_combine=="2326" ///
| marital_combine=="2362" | marital_combine=="6322" | marital_combine=="2366" ///
| marital_combine=="6326" | marital_combine=="6362" | marital_combine=="6366"
*marital = 2 (6) -> 3 -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="2323" | marital_combine=="2363" ///
| marital_combine=="6323" | marital_combine=="6363"
*marital = 2 (6) -> 3 -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="2332" | marital_combine=="2336" ///
| marital_combine=="6332" | marital_combine=="6336"
*marital = 2 (6) -> 3 -> 3 -> 3
replace marital_check=1 if marital_combine=="2333" | marital_combine=="6333"

```

[illegible]

```

*marital = 3 -> 2 (6) -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="3223" | marital_combine=="3226" ///
| marital_combine=="3262" | marital_combine=="3622" | marital_combine=="3266" ///
| marital_combine=="3626" | marital_combine=="3662" | marital_combine=="3666"
*marital = 3 -> 2 (6) -> 2 (6) -> 4
replace marital_check=1 if marital_combine=="3224" | marital_combine=="3264" ///
| marital_combine=="3624" | marital_combine=="3664"
*marital = 3 -> 2 (6) -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="3232" | marital_combine=="3236" ///
| marital_combine=="3632" | marital_combine=="3636"
*marital = 3 -> 2 (6) -> 3 -> 3
replace marital_check=1 if marital_combine=="3233" | marital_combine=="3633"
*marital = 3 -> 2 (6) -> 4 -> 2 (6)
replace marital_check=1 if marital_combine=="3242" | marital_combine=="3246" ///
| marital_combine=="3642" | marital_combine=="3646"
*marital = 3 -> 2 (6) -> 4 -> 3
replace marital_check=1 if marital_combine=="3243" | marital_combine=="3643"
*marital = 3 -> 2 (6) -> 4 -> 4
replace marital_check=1 if marital_combine=="3244" | marital_combine=="3644"
*marital = 3 -> 3 -> 2 (6) -> 2 (6)
replace marital_check=1 if marital_combine=="3322" | marital_combine=="3326" ///
| marital_combine=="3362" | marital_combine=="3366"
*marital = 3 -> 3 -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="3323" | marital_combine=="3363"
*marital = 3 -> 3 -> 2 (6) -> 4
replace marital_check=1 if marital_combine=="3324" | marital_combine=="3364"
*marital = 3 -> 3 -> 2 (6) -> 5
replace marital_check=1 if marital_combine=="3325" | marital_combine=="3365"
*marital = 3 -> 3 -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="3332" | marital_combine=="3336"
*marital = 3 -> 3 -> 3 -> 3
replace marital_check=1 if marital_combine=="3333"
*marital = 3 -> 4 -> 2 (6) -> 2 (6)
replace marital_check=1 if marital_combine=="3422" | marital_combine=="3426" ///
| marital_combine=="3462" | marital_combine=="3466"
*marital = 3 -> 4 -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="3423" | marital_combine=="3463"
*marital = 3 -> 4 -> 2 (6) -> 4
replace marital_check=1 if marital_combine=="3424" | marital_combine=="3464"
*marital = 3 -> 4 -> 2 (6) -> 5
replace marital_check=1 if marital_combine=="3425" | marital_combine=="3465"
*marital = 3 -> 4 -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="3432" | marital_combine=="3436"
*marital = 3 -> 4 -> 3 -> 3
replace marital_check=1 if marital_combine=="3433"
*marital = 3 -> 4 -> 4 -> 2 (6)
replace marital_check=1 if marital_combine=="3442" | marital_combine=="3446"
*marital = 3 -> 4 -> 4 -> 3
replace marital_check=1 if marital_combine=="3443"
*marital = 3 -> 4 -> 4 -> 4
replace marital_check=1 if marital_combine=="3444"
*marital = 3 -> 5 -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="3532" | marital_combine=="3536"
*marital = 3 -> 5 -> 3 -> 3
replace marital_check=1 if marital_combine=="3533"
*marital = 3 -> 5 -> 4 -> 2 (6)
replace marital_check=1 if marital_combine=="3542" | marital_combine=="3546"
*marital = 3 -> 5 -> 4 -> 3
replace marital_check=1 if marital_combine=="3543"
*marital = 3 -> 5 -> 4 -> 4
replace marital_check=1 if marital_combine=="3544"
*marital = 3 -> 5 -> 5 -> 3
replace marital_check=1 if marital_combine=="3553"
*marital = 3 -> 5 -> 5 -> 4
replace marital_check=1 if marital_combine=="3554"
*marital = 3 -> 5 -> 5 -> 5
replace marital_check=1 if marital_combine=="3555"

```

[illegible]

```

*marital = 5 -> 3 -> 2 (6) -> 4
replace marital_check=1 if marital_combine=="5324" | marital_combine=="5364"
*marital = 5 -> 3 -> 2 (6) -> 5
replace marital_check=1 if marital_combine=="5325" | marital_combine=="5365"
*marital = 5 -> 3 -> 3 -> 3
replace marital_check=1 if marital_combine=="5333"
*marital = 5 -> 3 -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="5332" | marital_combine=="5336"
*marital = 5 -> 4 -> 2 (6) -> 2 (6)
replace marital_check=1 if marital_combine=="5422" | marital_combine=="5426" ///
| marital_combine=="5462" | marital_combine=="5466"
*marital = 5 -> 4 -> 2 (6) -> 3
replace marital_check=1 if marital_combine=="5423" | marital_combine=="5426" ///
| marital_combine=="5462" | marital_combine=="5466"
*marital = 5 -> 4 -> 2 (6) -> 4
replace marital_check=1 if marital_combine=="5424" | marital_combine=="5464"
*marital = 5 -> 4 -> 2 (6) -> 5
replace marital_check=1 if marital_combine=="5425" | marital_combine=="5465"
*marital = 5 -> 4 -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="5432" | marital_combine=="5436"
*marital = 5 -> 4 -> 3 -> 3
replace marital_check=1 if marital_combine=="5433"
*marital = 5 -> 4 -> 4 -> 2 (6)
replace marital_check=1 if marital_combine=="5442" | marital_combine=="5446"
*marital = 5 -> 4 -> 4 -> 3
replace marital_check=1 if marital_combine=="5443"
*marital = 5 -> 4 -> 4 -> 4
replace marital_check=1 if marital_combine=="5444"
*marital = 5 -> 5 -> 3 -> 2 (6)
replace marital_check=1 if marital_combine=="5532" | marital_combine=="5536"
*marital = 5 -> 5 -> 3 -> 3
replace marital_check=1 if marital_combine=="5533"
*marital = 5 -> 5 -> 4 -> 2 (6)
replace marital_check=1 if marital_combine=="5542" | marital_combine=="5546"
*marital = 5 -> 5 -> 4 -> 3
replace marital_check=1 if marital_combine=="5543"
*marital = 5 -> 5 -> 4 -> 4
replace marital_check=1 if marital_combine=="5544"
*marital = 5 -> 5 -> 5 -> 3
replace marital_check=1 if marital_combine=="5553"
*marital = 5 -> 5 -> 5 -> 4
replace marital_check=1 if marital_combine=="5554"
*marital = 5 -> 5 -> 5 -> 5
replace marital_check=1 if marital_combine=="5555"

** replace marital_check **
replace marital_check=1 if dum_marital==0
replace marital_check=1 if marital_1=="."
replace marital_check=. if dum_marital==.
drop marital_1 marital_2 marital_3 marital_4 marital_combine
lab var marital_check "marital checking for the same person step3"

** 1.4 - gen dummy for the same individuals with marital status condition **
bysort caseid_new2: gen same_member14=1 if marital_check==1
replace same_member14=0 if same_member14==.
replace same_member14=. if marital_check==.
replace same_member14=0 if same_member13==0
lab def same_member14 1 "same" 0 "not same"
lab value same_member14 same_member14
lab var same_member14 "the same individual dummy step9 (marital status condition)"

**combine same_member in nonmatched data with matched data
gen same_member3=same_member14

```

```

** 1.5 - check number of unique individuals **
*count unique member step 5
by caseid_new, sort : gen count_memnum5=1 if _n==1 & same_member9==1
sort casehid caseid_new2 wave
lab var count_memnum5 "count unique of individual live in more than one wave step8"

*count unique member step 6
by caseid_new2, sort : gen count_memnum6=1 if _n==1 & same_member13==1
sort casehid caseid_new2 wave
lab var count_memnum6 "count unique of individual live in more than one wave step9"
*****

```

- สร้างเงื่อนไขเพื่อตรวจสอบข้อมูลตัวอย่างซ้ำรายบุคคล โดยใช้ caseid_new2 เมื่อกำหนดให้คนเดียวกันต้องมีเพศเดียวกัน และมีอายุห่างกันไม่เกิน 1 ปี นับจากการถูกสัมภาษณ์ในครั้งแรกจนถึงครั้งสุดท้าย นอกจากนี้ยังเพิ่มเงื่อนไขของจำนวนปีการศึกษาจะต้องเพิ่มขึ้นไม่เกิน 1 ปี นับจากการถูกสัมภาษณ์ในครั้งแรกจนถึงครั้งสุดท้าย และตั้งชื่อตัวแปรที่สอดคล้องกับเงื่อนไขนี้ว่า same_member15

```

***** STEP 10 check education level *****

** 1.1 - generate education change dummy **
/*browse casehid caseid_new wave round year sex age marital edulevel re_ed ///
same_member1 same_member2 same_member3 same_member4 number_waveid*/
sort casehid caseid_new2 wave
bysort caseid_new2 re_ed : egen dum_edu0=sum(count) if same_member13==1
gen dum_edu=number_waveid-dum_edu0
lab var number_waveid "count of wave for panel individuals"
sort casehid caseid_new2 wave
replace dum_edu=1 if dum_edu>0 & dum_edu!=.
lab def dum_edu 0 "not change" 1 "change"
lab value dum_edu dum_edu
lab var dum_edu "education level dummy"
drop dum_edu0

** 1.2 - generate year of schooling **
gen ysch=.

*code ysch 2001-2006 */
if year<2007 {
replace ysch = 0 if (edulevel == 0 | edulevel == 1 | edulevel == 2 | edulevel == 3)
replace ysch = 1 if (edulevel == 4)
replace ysch = 2 if (edulevel == 5 | edulevel == 84)
replace ysch = 3 if (edulevel == 6 | edulevel == 85)
replace ysch = 4 if (edulevel == 7 | edulevel == 86)
replace ysch = 5 if (edulevel == 8 | edulevel == 9)
replace ysch = 6 if (edulevel == 10 | edulevel == 35 | edulevel == 87)
replace ysch = 7 if (edulevel == 12 | edulevel == 37)
replace ysch = 8 if (edulevel == 13 | edulevel == 14 | edulevel == 38)
replace ysch = 9 if (edulevel == 15 | edulevel == 39 | edulevel == 88)
replace ysch = 10 if (edulevel == 17 | edulevel == 40 | edulevel == 89)
replace ysch = 11 if (edulevel == 18 | edulevel == 41)
replace ysch = 12 if (edulevel == 19 | edulevel == 42 | edulevel == 44 | ///
edulevel == 36 | edulevel == 58 | edulevel == 90)
replace ysch = 13 if (edulevel == 11 | edulevel == 21 | edulevel == 23 | ///
edulevel == 29 | edulevel == 78 | edulevel == 91)
replace ysch = 13 if (edulevel == 43 | edulevel == 46 | edulevel == 47)
replace ysch = 13 if (edulevel == 51 | edulevel == 60 | edulevel == 62 | ///
edulevel == 63)
}

```

```

replace ysch = 14 if (edulevel == 16 | edulevel == 22 | edulevel == 24 | ///
edulevel == 48 | edulevel == 45 | edulevel==59 | edulevel == 79 | edulevel == 92)
replace ysch = 14 if (edulevel == 49 | edulevel == 50 | edulevel == 52 | ///
edulevel == 61)
replace ysch = 14 if (edulevel == 64 | edulevel == 66)
replace ysch = 15 if (edulevel == 20 | edulevel == 25 | edulevel == 53)
replace ysch = 15 if (edulevel == 65 | edulevel == 67 | edulevel == 77 | ///
edulevel == 82 | edulevel == 80 | edulevel == 93)
replace ysch = 16 if (edulevel == 26 | edulevel == 30 | edulevel == 54 | ///
edulevel == 55 | edulevel == 81 | edulevel == 94)
replace ysch = 16 if (edulevel == 56 | edulevel == 57 | edulevel == 68)
replace ysch = 17 if (edulevel == 31)
replace ysch = 17 if (edulevel == 27)
replace ysch = 17 if (edulevel == 30) & (occuptn=="2141" | occuptn=="2224")
replace ysch = 18 if (edulevel == 32 | edulevel == 69)
replace ysch = 18 if (edulevel == 30) & (occuptn=="2221" | occuptn=="2222" ///
| occuptn=="2223")
replace ysch = 21 if (edulevel == 33 | edulevel == 70)
}
else if year>2006 {
/*Change to std. year of schooling 2007-2013*/
replace ysch = 0 if (GRADE_B == 0 | GRADE_B == 110 | GRADE_B == 111 | GRADE_B == 001)
replace ysch = 1 if GRADE_B== 211
replace ysch = 2 if GRADE_B == 212
replace ysch = 3 if GRADE_B == 213
replace ysch = 4 if (GRADE_B == 214 | GRADE_B == 219)
replace ysch = 5 if GRADE_B == 215
replace ysch = 6 if (GRADE_B == 216 | GRADE_B == 210 | GRADE_B == 240 | GRADE_B == 250)
replace ysch = 7 if (GRADE_B == 311 | GRADE_B == 321)
replace ysch = 8 if (GRADE_B == 312 | GRADE_B == 322 | GRADE_B == 319 | GRADE_B == 329)
replace ysch = 9 if (GRADE_B == 313 | GRADE_B == 323 | GRADE_B == 310 | GRADE_B == 320 ///
| GRADE_B == 330 | GRADE_B == 340 | GRADE_B == 350)
replace ysch = 10 if (GRADE_B == 411 | GRADE_B == 421)
replace ysch = 11 if (GRADE_B == 412 | GRADE_B == 422 | GRADE_B == 419 | GRADE_B == 429)
replace ysch = 12 if (GRADE_B == 413 | GRADE_B == 423 | GRADE_B == 420 | GRADE_B == 410 ///
| GRADE_B == 430 | GRADE_B == 440 | GRADE_B == 450)
replace ysch = 13 if (GRADE_B == 461 | GRADE_B == 511 | GRADE_B == 519 | GRADE_B == 521 ///
| GRADE_B == 611)
replace ysch = 14 if (GRADE_B == 462 | GRADE_B == 460 | GRADE_B == 510 | GRADE_B== 512 ///
| GRADE_B == 522 | GRADE_B == 529 | GRADE_B == 612 | GRADE_B == 619)
replace ysch = 15 if (GRADE_B == 520 | GRADE_B == 423 | GRADE_B == 613)
replace ysch = 16 if (GRADE_B == 610 | GRADE_B == 614 | GRADE_B == 630 | GRADE_B == 640 ///
| GRADE_B == 650)
replace ysch = 17 if (GRADE_B == 615 | GRADE_B == 660 | GRADE_B == 661 | GRADE_B == 711 ///
| GRADE_B == 719)
replace ysch = 18 if (GRADE_B == 616 | GRADE_B == 710 | GRADE_B == 712)
replace ysch = 19 if (GRADE_B == 760 | GRADE_B == 761 | GRADE_B == 811 | GRADE_B == 871)
replace ysch = 20 if (GRADE_B == 812 | GRADE_B == 819 | GRADE_B == 872)
replace ysch = 21 if (GRADE_B == 810 | GRADE_B == 873 | GRADE_B == 879)
replace ysch = 22 if (GRADE_B == 860 | GRADE_B == 861 | GRADE_B == 874)
replace ysch = 23 if (GRADE_B == 870 | GRADE_B == 875)
}
lab var ysch "Year of Schooling"

```

```

** 1.3 - generate dummy for increasing of year of schooling **
gen ysch_new=ysch
replace ysch_new=. if same_member2==0

bysort caseid: gen ysch0=ysch_new if _n==1 & same_member2==1
bysort caseid: egen ysch1=max(ysch0) if same_member2==1
drop ysch0
gen diff_ysch=ysch_new-ysch1
gen edu_increase0=1 if diff_ysch>=0 & diff_ysch<=1
replace edu_increase0=0 if diff_ysch<0 | diff_ysch>1
replace edu_increase0=1 if diff_ysch==. & same_member2==1
replace edu_increase0=. if same_member2==0
bysort caseid : egen edu_increase=min(edu_increase0) if same_member2==1
lab var edu_increase "checking of year of schooling does not decrease (0-1 year)"
drop ysch1 diff_ysch edu_increase0 ysch_new

** 1.4 - gen dummy for the same individuals with marital status not change **
bysort caseid_new2: gen same_member15=1 if edu_increase==1
replace same_member15=0 if same_member15==.
replace same_member15=. if same_member13==0
replace same_member15=0 if same_member13==0
lab def same_member15 1 "same" 0 "not same"
lab value same_member15 same_member15
lab var same_member15 "the same individual dummy step10 (check year of schooling)"
**combine same_member in nonmatched data with matched data
gen same_member4=same_member15

```

- สร้างเงื่อนไขเพื่อตรวจสอบข้อมูลตัวอย่างซ้ำรายบุคคล โดยใช้ caseid_new2 เมื่อกำหนดให้คนเดียวกันต้องมีเพศเดียวกัน และมีอายุห่างกันไม่เกิน 1 ปีนับจากการถูกสัมภาษณ์ในครั้งแรกจนถึงครั้งสุดท้าย นอกจากนี้ยังเพิ่มเงื่อนไขสถานภาพสมรส และจำนวนปีการศึกษาจะต้องเพิ่มขึ้นไม่เกิน 1 ปี นับจากการถูกสัมภาษณ์ในครั้งแรกจนถึงครั้งสุดท้าย และตั้งชื่อตัวแปรที่สอดคล้องกับเงื่อนไขนี้ว่า same_member16

```

** 1.5 - gen same individuals with marital and education condition
bysort caseid_new2: gen same_member16=1 if same_member14==1 & same_member15==1
replace same_member16=0 if same_member16==.
replace same_member16=. if same_member13==0
replace same_member16=0 if same_member13==0
lab def same_member16 1 "same" 0 "not same"
lab value same_member16 same_member16
lab var same_member16 "the same individual dummy step16 (check marital status and year of schooling)"
**combine same_member in nonmatched data with matched data
gen same_member5=same_member16

```

- เปลี่ยนตัวแปร same_member6 - same_member16 ที่ใช้ caseid_new และ caseid_new2 ให้เหมือนกับ same_member1- same_member5 ที่ใช้ caseid และระบุคำอธิบายพร้อมจัดลำดับของตัวแปร

```

**label variables
lab var same_member1 "the same individual dummy step1 (age 0-2)"
lab var same_member2 "the same individual dummy step2 (age 0-1)"
lab var same_member3 "the same individual dummy step3 (marital status condition)"
lab var same_member4 "the same individual dummy step4 (check year of schooling)"
lab var same_member5 "the same individual dummy step5 (check marital status and year of schooling)"

```

```

** order and sort
order casehid casehid caseid caseid_new caseid_new2
order wavehid wavehid2 waveid number_waveid count countw_samemem countwave_dum ///
countwave_idnew countwave_idnew2 count_memnum3 count_memnum4 count_memnum5 count_memnum6 ///
total_hhnum total_memnum samenumb dum_enum ///
firstenum lastenum dum_enum numb_members max_numb min_numb ///
labour head marital_check dum_marital dum_edu ysch edu_increase ///
same_member0 same_member1 same_member2 same_member3 same_member4 same_member5,after(enumdist)
sort casehid caseid_new2 wave

** label define **
lab def same_member1 1 "same" 0 "not same"
lab value same_member1 same_member1
lab value same_member2 same_member1
lab value same_member3 same_member1
lab value same_member4 same_member1
lab value same_member5 same_member1

** drop some variables
drop maxid same_member same_member6 same_member7 same_member8 same_member9 same_member10 ///
same_member11 same_member12 same_member13 same_member14 same_member15 same_member16

**re-gen number of total household and total member
drop total_hhnum total_memnum
by casehid,sort : gen total_hhnum=1 if _n==1
by caseid,sort : gen total_memnum=1 if _n==1
by caseid_new,sort : gen total_memnum1=1 if _n==1
by caseid_new2,sort : gen total_memnum2=1 if _n==1

lab var total_hhnum "marking for unique household (casehid)"
lab var total_memnum "marking for unique individuals (caseid)"
lab var total_memnum1 "marking for unique individuals (caseid_new)"
lab var total_memnum2 "marking for unique individuals (caseid_new2)"

**save as non-matched ORG data
save "${org}samsetA_rotation`m'_ORG_IDnonmatch.dta",replace
}

```

10. รวมข้อมูลที่ได้จากการใช้ caseid และ caseid_new ที่ใช้ในการเชื่อมโยงเพื่อค้นหาข้อมูลตัวอย่างซ้ำรายบุคคล แยกตาม rotation plan

[illegible]

```

** (1) append IDmatch and IDnonmatch **
forv m=1/4 {
  use "${org}samsetA_rotation`m'_ORG_IDmatch.dta", clear
  append using "${org}samsetA_rotation`m'_ORG_IDnonmatch.dta"
  order casehid caseid caseid_new caseid_new2
  order wavehid wavehid2 waveid number_waveid dumwaveid count countw_samemem ///
  countwave_dum ///
  countwave_idnew countwave_idnew2 count_memnum1 count_memnum2 count_memnum3 ///
  count_memnum4 count_memnum5 count_memnum6 ///
  total_hhnum total_memnum samenumb dum_enum ///
  firstenum lastenum dum_enum numb_members max_numb min_numb group ///
  diffage modesex increase labour head marital_check dum_marital dum_edu ysch edu_increase ///
  same_member0 same_member1 same_member2 same_member3 same_member4 same_member5,after(enumdist)

  *fill caseid_new and caseid_new2
  replace caseid_new=caseid if caseid_new=="
  replace caseid_new2=caseid if caseid_new2=="

  ** fill missing value of same_member with not the same person
  replace same_member3=0 if same_member3==.
  replace same_member4=0 if same_member4==.
  replace same_member5=0 if same_member5==.

  *sort
  sort casehid caseid wave

  *save as non-matched ORG data
  save "${org}samsetA_rotation`m'_ORG_Combine.dta",replace
}
*-----*

** (2) append ORG data for samset A **
use "${org}samsetA_rotation1_ORG_Combine.dta",clear
forv m=2/4 {
  append using "${org}samsetA_rotation`m'_ORG_Combine.dta"
}

** checking number_waveid and replacing to not same individuals panel (live in one wave)
drop number_waveid
bysort caseid : egen number_waveid=sum(count)
bysort caseid_new : egen number_waveidnew=sum(count)
bysort caseid_new2 : egen number_waveidnew2=sum(count)
replace same_member1=0 if number_waveidnew==1
replace same_member2=0 if number_waveidnew==1
replace same_member3=0 if number_waveidnew==1
replace same_member4=0 if number_waveidnew==1
replace same_member5=0 if number_waveidnew==1

**label variables
lab var number_waveid "number of wave baseon caseid"
lab var number_waveidnew "number of wave baseon caseid_new"
lab var number_waveidnew2 "number of wave baseon caseid_new2"

**re-gen number of total household and total member
drop total_hhnum total_memnum total_memnum1 total_memnum2 numhh
by casehid,sort : gen total_hhnum=1 if _n==1
by caseid,sort : gen total_memnum=1 if _n==1
by caseid_new,sort : gen total_memnum1=1 if _n==1
by caseid_new2,sort : gen total_memnum2=1 if _n==1

lab var total_hhnum "marking for unique household (casehid)"
lab var total_memnum "marking for unique individuals (caseid)"
lab var total_memnum1 "marking for unique individuals (caseid_new)"
lab var total_memnum2 "marking for unique individuals (caseid_new2)"

```

```

** count same member with same age, sex and household
bysort casehid wave sex age : egen sameage_group0=sum(count)
replace sameage_group0=0 if sameage_group0==1
replace sameage_group0=1 if sameage_group0>1
bysort casehid : egen sameage_group=max(sameage_group0)
drop sameage_group0
lab def sameage_group 0 "not same" 1 "same"
lab value sameage_group sameage_group
lab var sameage_group "Checking member in the same household with same age and sex"

** same as ORG data
save "{org}samsetA_ORG.dta",replace
*-----*

** (3) erase data **
forv m=1/4 {
  erase "{org}samsetA_rotation`m'_ORG_IDmatch.dta"
  erase "{org}samsetA_rotation`m'_ORG_IDnonmatch.dta"
  erase "{org}samsetA_rotation`m'_ORG_Combine.dta"
}
*#####

```

ภาคผนวก ค.

สรุปรายงานการประชุมกับผู้เชี่ยวชาญจากสำนักงานสถิติแห่งชาติ

รายงานการประชุมหารือ

เรื่องแผนการเลือกตัวอย่างสำหรับโครงการสำรวจภาวะการทำงาน ครั้งที่1/2560

เมื่อวันพฤหัสบดีที่ 12 มกราคม 2560

ณ ห้องประชุม ชั้น 3 สำนักนโยบายและวิชาการสถิติ สำนักงานสถิติแห่งชาติ

ผู้เข้าร่วมประชุม

1. นางธิดา ชัยเดชอัครกุล ผอ.กลุ่มสถิติแรงงาน สำนักงานสถิติแห่งชาติ
2. นายอานนท์ จันทวิช ผู้ตรวจราชการกรม สำนักงานสถิติแห่งชาติ
3. นางหทัยชนก พรระเจริญ ผอ.กลุ่มระเบียบวิธีสถิติ สำนักงานสถิติแห่งชาติ
4. นางสาวพรพรรณ แก้วศรีงาม ผอ.กลุ่มคลังข้อมูล สำนักงานสถิติแห่งชาติ
5. นางสาวบงกช วิบูลย์ธนานันต์ ผอ.สำนักนโยบายและวิชาการสถิติ สำนักงานสถิติแห่งชาติ
6. นางสาววรัญญา สุขวงศ์ นักวิชาการสถิติ ชำนาญการ สำนักงานสถิติแห่งชาติ
7. ดร.ภัทรพรรณ อดทน หัวหน้าโครงการวิจัย นักวิจัย สถาบันวิจัยเพื่อการประเมินและ
ออกแบบนโยบาย ม.หอการค้าไทย
8. นางสาววาสนา จันทรรณ นักวิจัย สถาบันวิจัยเพื่อการประเมินและออกแบบนโยบาย
ม.หอการค้าไทย

เริ่มประชุมเวลา 13.30 น.

วาระที่ 1. เรื่องที่เสนอให้ที่ประชุมทราบ

1.1 โครงการ “การจัดทำฐานข้อมูลแบบตัวอย่างซ้ำจากข้อมูลภาวะการทำงานของประชากร”
ได้รับทุนสนับสนุนจาก สกว. โดยมีระยะเวลาดำเนินการ 1 ปี นับตั้งแต่วันที่ 3 ตุลาคม 2559 ถึง 2
ตุลาคม 2560

มติ ที่ประชุมรับทราบ

1.2 วัตถุประสงค์ โครงการ “การจัดทำฐานข้อมูลแบบตัวอย่างซ้ำจากข้อมูลภาวะการทำงานของประชากร”

- เพื่อจัดทำฐานข้อมูลภาวะการทำงานของประชากรแบบตัวอย่างซ้ำจากข้อมูลการสำรวจแบบ cross-sectional ในแต่ละไตรมาส ตามแผนแบบการเลือกตัวอย่าง 2-2-2 สำหรับโครงการสำรวจภาวะการทำงานของประชากร
- เพื่อส่งเสริมให้นักวิจัยได้ใช้ประโยชน์จากวิธีการจัดทำฐานข้อมูลรายบุคคลแบบตัวอย่างซ้ำ มติ ที่ประชุมรับทราบ

วาระที่ 2. เรื่องที่เสนอให้ที่ประชุมพิจารณา

2.1 แผนแบบการเลือกตัวอย่าง 2-2-2 ที่นักวิจัยได้ทำการศึกษาตั้งแต่ปี 2545-2556 จากข้อมูล LFS สอดคล้องกับแผนแบบการเลือกตัวอย่าง 2-2-2 ของสำนักงานสถิติแห่งชาติหรือไม่ อย่างไร?

มติ ที่ประชุมชี้แจงว่า แผนแบบการเลือกตัวอย่าง 2-2-2 ของสำนักงานสถิติมีการเปลี่ยนแผนแบบ โดยแบ่งเป็น 2 ช่วง คือ ช่วงระหว่างปี 2545-2554 ซึ่งเป็น Rotation ชุดเดิม และอีกช่วงคือ ระหว่างปี 2555-2556 จะเป็น Rotation ชุดใหม่ โดยนักวิจัยสามารถขอเอกสารแผนการเลือกตัวอย่างจากสำนักงานสถิติแห่งชาติได้โดยตรง

2.2 ฐานข้อมูลที่กำลังจะจัดทำขึ้นนี้ นักวิจัยสามารถทำ MOU ร่วมกับสำนักงานสถิติเพื่อเผยแพร่ได้หรือไม่ อย่างไร?

มติ ที่ประชุมชี้แจงว่า ทางสำนักงานสถิติไม่สามารถให้นักวิจัยเผยแพร่ฐานข้อมูลดังกล่าวนี้ได้ เนื่องจากพระราชบัญญัติว่าด้วยการกระทำผิดเกี่ยวกับคอมพิวเตอร์ พ.ศ. 2550 มีบทลงโทษระบุไว้ชัดเจน ทำให้ผู้ปฏิบัติหน้าที่ที่เกี่ยวข้องทั้งจากสำนักงานสถิติแห่งชาติและนักวิจัยในโครงการมีความผิด

2.3 นักวิจัยสามารถเผยแพร่โค้ดคำสั่งวิธีการจัดทำฐานข้อมูลตามแผนแบบการเลือกตัวอย่าง 2-2-2 ได้หรือไม่ อย่างไร?

มติ ที่ประชุมเห็นควรว่า นักวิจัยสามารถเผยแพร่โค้ดคำสั่งในการจัดการข้อมูลได้ แต่ควรอธิบายถึงข้อควรระวังและลักษณะของข้อมูลที่นักวิจัยจะนำไปใช้ โดยนักวิจัยจะนำไปปรับแก้และเขียนคำอธิบายไว้ในคู่มือการใช้โค้ดคำสั่ง

2.4 ข้อมูลการสำรวจในแต่ละไตรมาสมีการถามใหม่ในทุกไตรมาสหรือไม่? และตัวแปรรหัสครัวเรือนและสมาชิกครัวเรือนมีการเปลี่ยนแปลงทุกไตรมาสหรือไม่? อย่างไร?

มติ ที่ประชุมชี้แจงว่า มีการถามข้อมูลใหม่ในทุกไตรมาส (ไม่ได้มีการจัดทำ roster) และอาจทำให้ข้อมูลเกิด non-sampling error นั่นคือ ข้อมูลบางอย่างของแต่ละคนเกิดความคลาดเคลื่อนและอาจจะไม่ตรงกันในแต่ละไตรมาส เช่น อายุ, ปี พ.ศ. ที่เกิด, อาชีพ, ระดับการศึกษา เป็นต้น

ส่วนรหัสครัวเรือนและรหัสสมาชิกครัวเรือนยังเป็นหมายเลขเดิม หากทีมสัมภาษณ์พบว่า ครัวเรือนยังอยู่ที่เดิมและสามารถให้ข้อมูลแก่ทีมสัมภาษณ์ได้ แต่หากครัวเรือนดังกล่าวย้ายออกและไม่สามารถตามตัวได้ ครัวเรือนนั้นจะถูกทดแทนด้วยครัวเรือนใหม่ซึ่งยังคงใช้เลขของครัวเรือนเดิม ดังนั้น นักวิจัยต้องตรวจสอบก่อนว่า ครัวเรือนนั้นถูกทดแทนด้วยครัวเรือนอื่นหรือไม่ (ดูตัวแปร enumtype ผลการแจกแจงประกอบ) เพื่อใช้ยืนยันว่าครัวเรือนดังกล่าวเป็นครัวเรือนเดียวกัน

รหัสสมาชิกครัวเรือน จะไล่ตามความสัมพันธ์กับหัวหน้าครัวเรือน ซึ่งในแต่ละไตรมาสอาจมีการเปลี่ยนแปลงหัวหน้าครัวเรือน ซึ่งทำให้รหัสสมาชิกครัวเรือนมีการรันใหม่ทุกครั้ง ดังนั้น นักวิจัยจะต้องระมัดระวังในการไต่ตรหัสสมาชิกครัวเรือนคนดังกล่าวให้เหมือนกันในทุกไตรมาส

2.5 ตัวแปร Identification ประกอบไปด้วยตัวแปรอะไรบ้าง เพื่อใช้ศึกษาแผนแบบการเลือกตัวอย่างแบบ 2-2-2

มติ ที่ประชุมโดยผู้เชี่ยวชาญจากสำนักงานสถิติแห่งชาติ ระบุว่า ID ระดับครัวเรือนในแต่ละไตรมาสของแต่ละปี ประกอบไปด้วย ตัวแปร region, province, area, psunumber, samset, hhnumber และ ID ระดับบุคคลในแต่ละไตรมาสของแต่ละปี ประกอบไปด้วย ตัวแปร region, province, area, psunumber, samset, hhnumber และ memnumber ตามลำดับ

2.6 นักวิจัยจะทราบได้หรือไม่ว่าบุคคลนั้นเป็นผู้ตอบคำถามแทนสมาชิกคนอื่นๆ ในครัวเรือน

มติ ที่ประชุมชี้แจงว่า มีตัวแปรที่ระบุว่าใครเป็นผู้ตอบแบบสอบถามแทนบุคคลอื่นๆ ในครัวเรือน โดยในเบื้องต้นนักวิจัยจะกลับไปทำการศึกษาข้อมูลอีกครั้ง และหากนักวิจัยพบว่าข้อมูลไม่มีตัวแปรในส่วนดังกล่าว นักวิจัยสามารถสอบถามไปยังผู้เชี่ยวชาญเพื่อขอข้อมูลในส่วนนี้ได้

วาระที่ 3. เรื่องอื่น ๆ (ถ้ามี)

3.1 ทางผู้เชี่ยวชาญให้ข้อมูลกับนักวิจัยเรื่องการใช้แท็บเล็ตในการทำสำรวจภาวะการทำงานของประชากร โดยทางสำนักงานสถิติแห่งชาติเริ่มใช้ในปีแรกคือ พ.ศ. 2557

มติ ที่ประชุมแจ้งให้นักวิจัยศึกษาเพิ่มเติม

3.2 นักวิจัยควรตรวจสอบร้อยละของครัวเรือนตัวอย่างซ้ำในไตรมาสที่ติดกัน สำหรับแผนการเลือกตัวอย่าง ในช่วงปี 2545-2554 ซึ่งจะคิดเป็นร้อยละ 25, 50 และ 75 ตามลำดับ โดยอาจมีจำนวนร้อยละไม่ตรงกันในบางปี

มติ ที่ประชุมแจ้งให้นักวิจัยได้รับทราบ

3.3 ผู้เชี่ยวชาญเห็นว่า นักวิจัยควรเสนอวิธีการของแผนแบบการเลือกตัวอย่างซ้ำแบบ 2-2-2 ของสำนักงานสถิติแห่งชาติเทียบกับแผนการเลือกตัวอย่างซ้ำของ CPS สหรัฐอเมริกา

มติ ที่ประชุมแจ้งให้นักวิจัยเสนอผลในรายงานของโครงการ

3.4 ตรวจสอบประเทศในกลุ่มอาเซียนว่ามีประเทศใดที่ใช้แผนการเลือกตัวอย่างจาก ORG plan อีกหรือไม่ นอกจากประเทศไทย เช่น เวียดนาม เป็นต้น

มติ ที่ประชุมแจ้งให้นักวิจัยศึกษาเพิ่มเติม

3.5 นักวิจัยควรเสนอขั้นตอนวิธีการจัดการข้อมูลอย่างละเอียดไว้ในคู่มือการใช้งาน

มติ ที่ประชุมแจ้งให้นักวิจัยศึกษาเพิ่มเติม

เลิกประชุมเวลา 15.30 น.

รายงานการประชุมเพื่อให้ข้อคิดเห็นและข้อเสนอแนะ
ภายใต้โครงการการจัดทำฐานข้อมูลจากข้อมูลภาวะการทำงานของประชากร
ครั้งที่ 2/2560
เมื่อวันพฤหัสบดีที่ 9 พฤศจิกายน 2560
ณ ห้องประชุมกองเศรษฐกิจ ชั้น 3 สำนักงานสถิติแห่งชาติ

ผู้เข้าร่วมประชุม

1. นางธิดา ชัยเดชอัครกุล ผอ.กลุ่มสถิติแรงงาน สำนักงานสถิติแห่งชาติ
2. นายอานนท์ จันทวิช ผู้ตรวจราชการ สำนักงานสถิติแห่งชาติ
3. นางน้ำผึ้ง เชิดชูพงษ์ ผอ.สำนักสถิติเศรษฐกิจและสังคม สำนักงานสถิติแห่งชาติ
4. นางสาวพรพรรณ แก้วศรีงาม ผอ.กลุ่มคลังข้อมูล สำนักงานสถิติแห่งชาติ
5. นางสาวบงกช วิบูลย์ธนานันต์ ผอ.สำนักนโยบายและวิชาการสถิติ สำนักงานสถิติแห่งชาติ
6. นางสาววรัญญา สุขวงศ์ นักวิชาการสถิติชำนาญการ สำนักงานสถิติแห่งชาติ
7. ดร.ภัทรพรรณ อดทน หัวหน้าโครงการวิจัย นักวิจัย สถาบันวิจัยเพื่อการประเมินและ
ออกแบบนโยบาย ม.หอการค้าไทย
8. นางสาววาสนิณี จันทรร นักวิจัย สถาบันวิจัยเพื่อการประเมินและออกแบบนโยบาย
ม.หอการค้าไทย

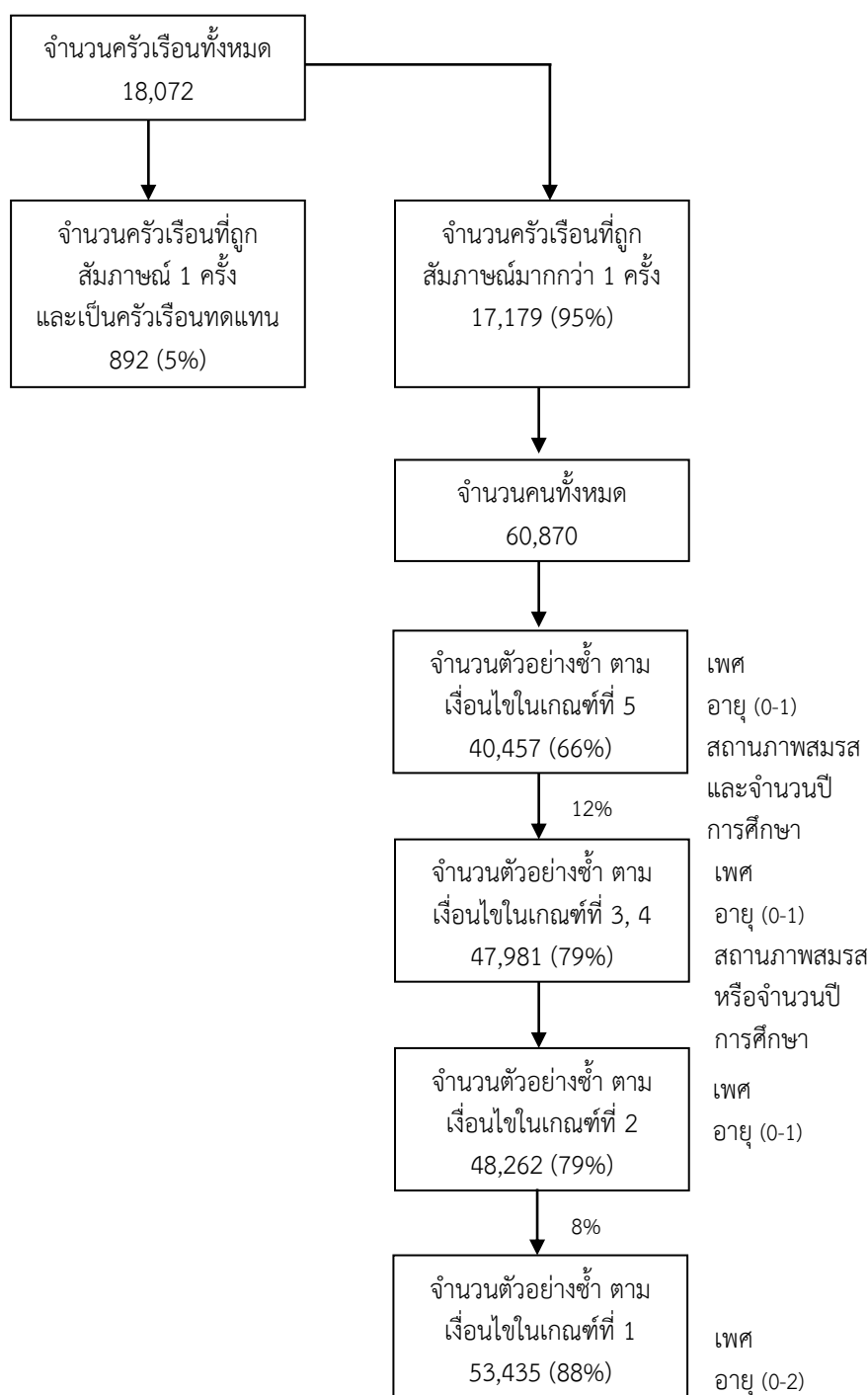
เริ่มประชุมเวลา 13.30 น.

วาระที่ 1. เรื่องที่เสนอให้ที่ประชุมทราบ

1.1 ความก้าวหน้าในการดำเนินงานของโครงการฯ นับตั้งแต่วันที่ 3 ตุลาคม 2559 ถึง 2 ตุลาคม 2560 โดยโครงการได้จัดทำคำศัพท์เพื่อใช้ในการจัดการข้อมูลตัวอย่างซ้ำรายบุคคลจากข้อมูลภาวะการทำงานของประชากร และมีเกณฑ์ที่นักวิจัยใช้ในการค้นหาข้อมูลตัวอย่างซ้ำรายบุคคล ซึ่งแบ่งออกเป็น 5 เกณฑ์ ดังนี้

- 1) กลุ่มตัวอย่างซ้ำคนเดียวกันต้องมีเพศเดียวกัน และอายุไม่เกิน 2 ปี ตั้งแต่การถูกสัมภาษณ์ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่สัมภาษณ์
- 2) กลุ่มตัวอย่างซ้ำคนเดียวกันต้องมีเพศเดียวกัน และอายุไม่เกิน 1 ปี ตั้งแต่การถูกสัมภาษณ์ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่สัมภาษณ์
- 3) กลุ่มตัวอย่างซ้ำคนเดียวกันต้องมีเพศเดียวกัน และอายุไม่เกิน 1 ปี และสถานภาพสมรสต้องเป็นไปตามเงื่อนไขของแต่ละสถานภาพ
- 4) กลุ่มตัวอย่างซ้ำคนเดียวกันต้องมีเพศเดียวกัน, อายุไม่เกิน 1 ปี และจำนวนปีการศึกษาจะต้องมีจำนวนเท่าเดิมหรือเพิ่มขึ้นไม่เกิน 1 ปีการศึกษานับตั้งแต่การถูกสัมภาษณ์ครั้งแรกจนถึงครั้งสุดท้าย
- 5) กลุ่มตัวอย่างซ้ำคนเดียวกันต้องมีเพศเดียวกัน, อายุไม่เกิน 1 ปี, สถานภาพสมรสเป็นไปตามเงื่อนไข และจำนวนปีการศึกษาจะต้องมีจำนวนเท่าเดิมหรือเพิ่มขึ้นไม่เกิน 1 ปีการศึกษานับตั้งแต่การถูกสัมภาษณ์ครั้งแรกจนถึงครั้งสุดท้าย

จากเกณฑ์ข้างต้น นักวิจัยพบว่า จำนวนข้อมูลกลุ่มตัวอย่างซ้ำรายบุคคลเฉลี่ยต่อ 1 rotation plan สามารถสรุปได้ดังนี้



มติ ที่ประชุมรับทราบ และชี้แจงเพิ่มเติมว่า นักวิจัยควรแสดงจำนวนและร้อยละของข้อมูลตัวอย่างซ้ำ รายบุคคลที่ถูกสัมภาษณ์ครบทุกครั้งด้วย นอกจากนี้ ผู้เชี่ยวชาญได้ให้ข้อสังเกตว่าการจัดทำข้อมูลตัวอย่างซ้ำ ในครั้งนี้จะเป็นประโยชน์อย่างยิ่งต่อเจ้าหน้าที่ที่ดูแลฐานข้อมูลภาวการณ์ทำงานของประชากร ซึ่งช่วยในการ ตรวจสอบความถูกต้องของข้อมูลและการขาดหายไปของกลุ่มตัวอย่างตามแผนการสุ่มตัวอย่าง

วาระที่ 2. เรื่องที่เสนอให้ที่ประชุมพิจารณา

2.1 ข้อเสนอแนะเพิ่มเติมเกี่ยวกับระเบียบวิธีวิจัย

มติ ที่ประชุมเห็นด้วยกับเกณฑ์การจัดทำฐานข้อมูลตัวอย่างซ้ำรายบุคคลทั้ง 5 เกณฑ์ ซึ่งได้ใช้ตัวแปรที่คาดว่า สามารถนำมาค้นหาตัวอย่างซ้ำรายบุคคลได้ จากทั้งตัวแปรเพศ อายุ สถานภาพสมรส และระดับการศึกษา ซึ่งตัวแปรอื่น ๆ ก็ไม่มีความเหมาะสมในการนำมาพิจารณา นอกจากนี้ นักวิจัยได้ปรึกษาถึงประเด็นการใช้ค่าถ่วงน้ำหนักของประชากร (weight of population) สำหรับข้อมูลตัวอย่างซ้ำรายบุคคล ซึ่งนักวิจัยเองไม่สามารถคำนวณใหม่ได้ (re-calculate) เนื่องจากไม่ทราบจำนวนประชากรในแต่ละ rotation sampling ดังนั้น ทางผู้เชี่ยวชาญจากสำนักงานสถิติได้ชี้แจงว่า นักวิจัยสามารถทำเรื่องขอทราบจำนวนประชากรในแต่ละ rotation sampling ได้กับเจ้าหน้าที่โครงการโดยตรง

2.2 การเผยแพร่ข้อมูล/ชุดคำสั่ง ซึ่งนักวิจัยได้เสนอ 2 แนวทาง คือ

- 1) เผยแพร่ข้อมูลผ่านการดำเนินการจัดทำเป็น ORG เรียบร้อยแล้ว
- 2) เผยแพร่ matching id

มติ ที่ประชุมสรุปว่า นักวิจัยสามารถเผยแพร่ matching id ผ่านเว็บไซต์ของทางสำนักงานสถิติแห่งชาติได้ พร้อมทั้งนักวิจัยจะจัดการเผยแพร่ชุดคำสั่งสำหรับการจัดการฐานข้อมูล

2.3 การจัดอบรมการใช้ข้อมูลและชุดคำสั่ง โดยโครงการฯ จะจัดอบรมให้กับนักวิจัย นิสิต นักศึกษา ผู้ที่สนใจใช้ข้อมูลตัวอย่างซ้ำรายบุคคลซึ่งคาดว่าจะจัดในช่วงเดือนมกราคม ปี 2561 และจะทำหนังสือเรียนเชิญผู้เชี่ยวชาญจากสำนักงานสถิติแห่งชาติ

มติ ที่ประชุมรับทราบ

เลิกประชุมเวลา 15.00 น.

รายงานการอบรมการใช้ฐานข้อมูลแบบตัวอย่างซ้ำจากข้อมูลภาวะการทำงานของประชากร

ณ ห้องประชุมศูนย์วิจัยมหาวิทยาลัยชิคาโก-มหาวิทยาลัยหอการค้าไทย

วันอังคารที่ 30 มกราคม 2561 เวลา 9.30 -11.30 น.

ผู้เข้าร่วมประชุม

1. ดร.ศศิวิมล วรณศิริ ปรีณวัฒน์ คณะเศรษฐศาสตร์ มหาวิทยาลัยหอการค้าไทย
2. ดร.เนื่อแพร เล็กเฟื่องฟู คณะเศรษฐศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
3. ดร.วรประภา นาควัชร คณะเศรษฐศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
4. คุณจุมล กุลโท นิสิตปริญญาโท คณะเศรษฐศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
5. ดร.มณเฑียร สติมานนท์ คณะเศรษฐศาสตร์ มหาวิทยาลัยธรรมศาสตร์
6. คุณกิตติภูมิ วิเศษศักดิ์ นักวิจัยการระดับชำนาญการพิเศษ
สถาบันเสริมศึกษาและทรัพยากร มหาวิทยาลัยธรรมศาสตร์
7. ดร.ยุรนนท์ ตามกาล นักวิจัยระดับปฏิบัติการ
สถาบันเสริมศึกษาและทรัพยากร มหาวิทยาลัยธรรมศาสตร์
8. คุณอุไรวรรณ รุ่งไธัญ นักวิจัยระดับปฏิบัติการ
สถาบันเสริมศึกษาและทรัพยากร มหาวิทยาลัยธรรมศาสตร์
9. คุณศรีษฐา อ่อนแก้ว นักวิจัยระดับปฏิบัติการ
สถาบันเสริมศึกษาและทรัพยากร มหาวิทยาลัยธรรมศาสตร์
10. คุณอานนท์ จันทวิช ผู้ตรวจราชการ สำนักงานสถิติแห่งชาติ
11. คุณปฐมวดี บรรจุทรัพย์ นักศึกษาปริญญาโท คณะเศรษฐศาสตร์ มหาวิทยาลัยธรรมศาสตร์
12. คุณสัจจา ดวงชัยอยู่สุข นักศึกษาปริญญาเอก มหาวิทยาลัยหอการค้าไทย
13. คุณวรุตม์ สามารถ นักศึกษาปริญญาเอก มหาวิทยาลัยหอการค้าไทย
14. คุณวิษวรัต ชูจันทร์ นักศึกษาปริญญาเอก มหาวิทยาลัยหอการค้าไทย
15. คุณพร้อมพงษ์ แสงขำ นักศึกษาปริญญาเอก มหาวิทยาลัยหอการค้าไทย
16. คุณปิยนุช ภูพัฒน์สกุล นักศึกษาปริญญาเอก มหาวิทยาลัยหอการค้าไทย
17. คุณโสธยา แพสุวรรณ นักวิจัย สถาบันวิจัยเพื่อการประเมินและออกแบบนโยบาย
18. คุณบุญสิทธิ์ สุทธินุ่น นักวิชาการสถิติชำนาญการ สำนักงานสถิติแห่งชาติ

19. คุณพัชยา เลาสุทแสง นักวิชาการสถิติชำนาญการ สำนักงานสถิติแห่งชาติ
20. คุณพรชนันชฌมน แจ้งถิ่นป่า นักวิชาการสถิติปฏิบัติการสำนักงานสถิติแห่งชาติ
21. ดร.วีระชาติ กิเลนทอง ผู้อำนวยการ สถาบันวิจัยเพื่อการประเมินและออกแบบนโยบาย
22. ดร.ภัทรพรรณ อดทน นักวิจัย สถาบันวิจัยเพื่อการประเมินและออกแบบนโยบาย
23. คุณไพรุส อับดุลเลาะห์ นักวิจัย สถาบันวิจัยเพื่อการประเมินและออกแบบนโยบาย
24. คุณวาสิณี จันทร์ธร นักวิจัย สถาบันวิจัยเพื่อการประเมินและออกแบบนโยบาย

กำหนดการประชุม

- | | |
|----------|--|
| 9:00 am | Registration |
| 9:30 am | Introduction to Constructing A Panel Data from the Labor Force Survey of Thailand
(ดร.ภัทรพรรณ อดทน สถาบันวิจัยเพื่อการประเมินและออกแบบนโยบาย.) |
| 10.00am | How to use Constructing A Panel Data from the Labor Force Survey of Thailand
(ดร.ภัทรพรรณ อดทน สถาบันวิจัยเพื่อการประเมินและออกแบบนโยบาย.) |
| 11.00 am | Questions and Answers |
| 11.30am | Closing |

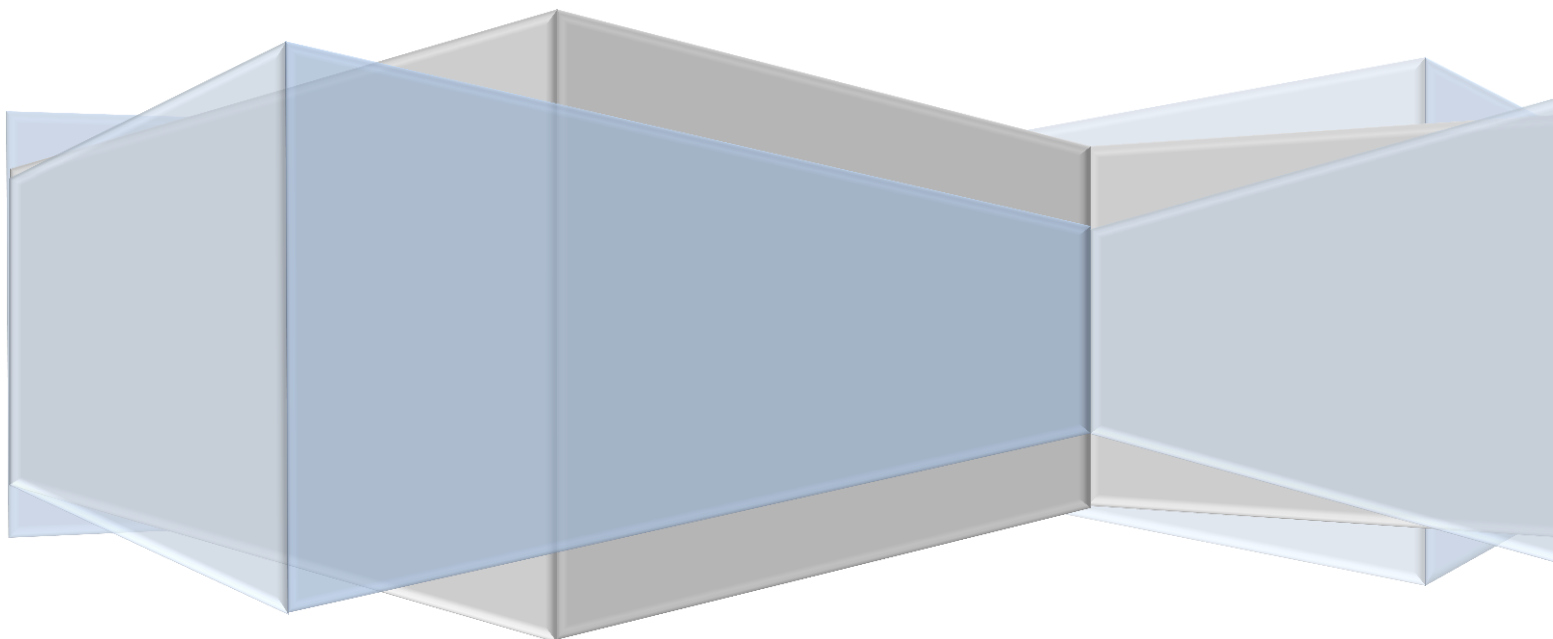
วิธีการสร้างและเผยแพร่ข้อมูลแบบตัวอย่างซ้ำจากข้อมูลภาวะการทำงานของประชากร

1. การใช้ Matching ID
 - ของข้อมูล LFS จากสำนักงานสถิติแห่งชาติ
 - จัดเตรียมข้อมูล LFS ให้อยู่ในรูปของ Stata file
 - ดาวน์โหลดข้อมูล Matching ID และ Stata commands
 - ใช้ Stata commands เพื่อเชื่อมโยงข้อมูล LFS และ Matching
2. การใช้ Stata commands
 - ของข้อมูล LFS จากสำนักงานสถิติแห่งชาติ
 - ดาวน์โหลด Stata commands
 - ใช้ Stata commands ในการสร้างข้อมูล ORG
3. การใช้บริการจัดการข้อมูลจากสถาบันวิจัยเพื่อการประเมินและออกแบบนโยบาย



UTCC
Research Institute for
Policy Evaluation & Design: RIPED
สถาบันวิจัยเพื่อการประเมินและออกแบบนโยบาย

คู่มือการใช้งาน
สำหรับการจัดการฐานข้อมูลตัวอย่างซ้ำ (ORG)
จากข้อมูลภาวการณ์ทำงานของประชากร
โดย
สถาบันวิจัยเพื่อการประเมินและออกแบบนโยบาย
มหาวิทยาลัยหอการค้าไทย



คำนำ

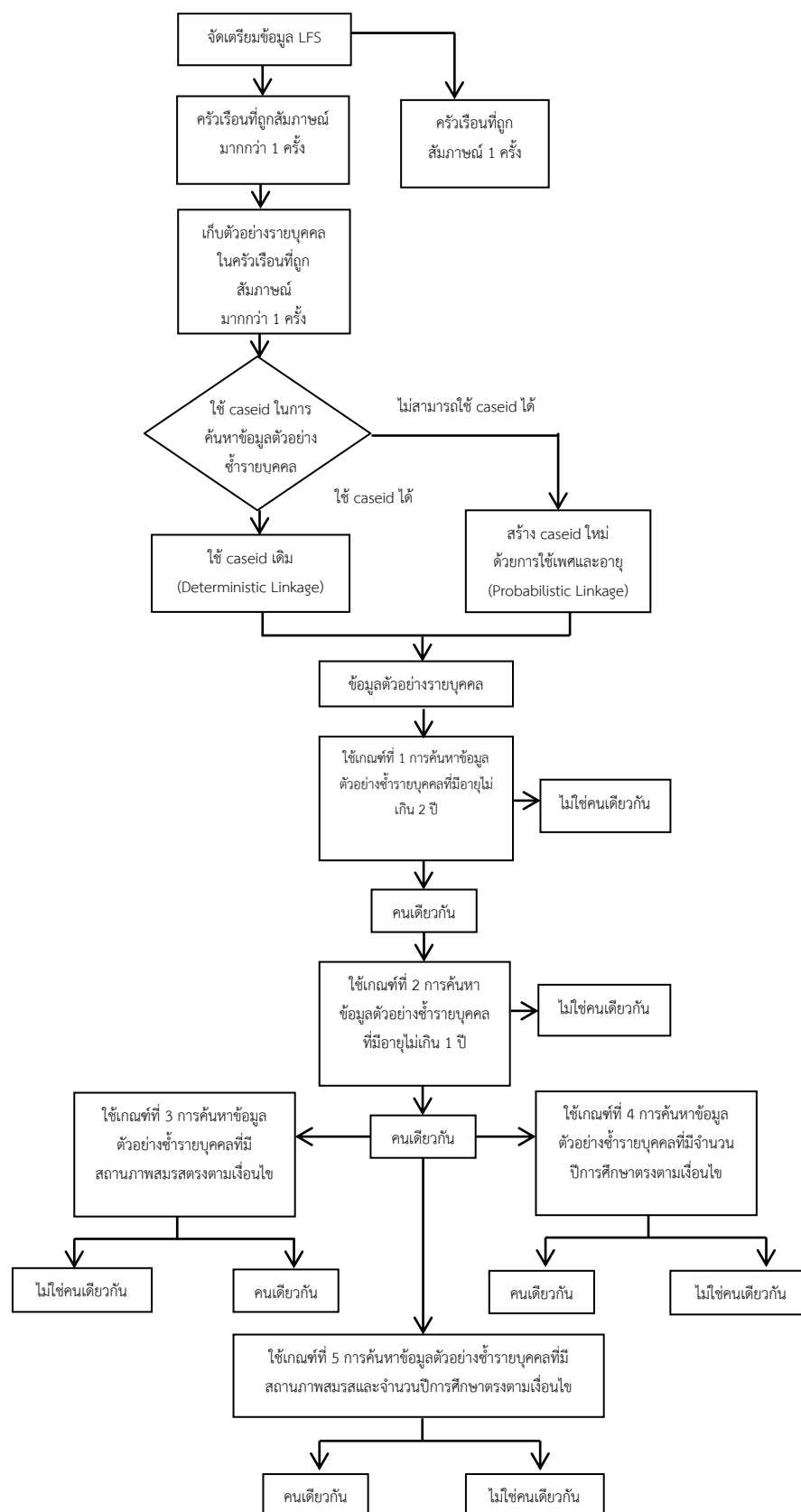
คู่มือการใช้งานชุดคำสั่งจัดทำขึ้นภายใต้โครงการ “การจัดทำฐานข้อมูลแบบตัวอย่างซ้ำจากข้อมูลภาวะการทำงานของประชากร” ซึ่งได้รับทุนจากสำนักงานกองทุนสนับสนุนการวิจัย (สกว.) โดยโครงการจะนำข้อมูลการสำรวจภาวะการทำงานของประชากร (Labor Force Survey: LFS) ของสำนักงานสถิติแห่งชาติ นับตั้งแต่ปี พ.ศ. 2545 – 2556 มาสร้างข้อมูลตัวอย่างซ้ำด้านแรงงานของประชากรไทยในส่วนของข้อมูล ORG โดยจะเผยแพร่ Matching ID และ STATA code พร้อมคู่มือการสร้างข้อมูลตัวอย่างซ้ำในเวปไซต์ เพื่อช่วยให้นักวิจัยที่มีข้อมูลการสำรวจภาวะการทำงานของประชากรสามารถสร้างข้อมูลได้อย่างสะดวกมากขึ้น

คณะนักวิจัย

สารบัญ

เรื่อง	หน้า
คำนำ	1
1. กระบวนการจัดทำข้อมูล ORG	3
2. วิธีที่ 1 ขั้นตอนการใช้ Matching ID เพื่อสร้างข้อมูล ORG	4
3. วิธีที่ 2 ขั้นตอนการใช้โค้ดคำสั่ง เพื่อสร้างข้อมูล ORG	7
4. เกณฑ์ที่ใช้ในการค้นหาข้อมูลตัวอย่างข้าราชการบุคคล	17

1. กระบวนการจัดทำข้อมูล ORG



2. วิธีที่ 1 ขั้นตอนการใช้ Matching ID เพื่อสร้างข้อมูล ORG

1. เตรียมข้อมูลภาวะการทำงานของประชากรจากสำนักงานสถิติแห่งชาติ ตั้งแต่ปี 2002-2013 (พ.ศ. 2545-2556) ซึ่งมีทั้งไฟล์ในรูปแบบ .sav (SPSS file) และ .dat (Text file)
2. บันทึกข้อมูลภาวะการทำงานของประชากรที่อยู่ในรูปของไฟล์ .dta (Stata file) ให้อยู่ในโฟลเดอร์ที่ชื่อ data

Name	Date modified	Type	Size
 lfs2002q1	25/1/2561 10:03	Stata Dataset	2,182 KB
 lfs2002q2	25/1/2561 10:03	Stata Dataset	21 KB
 lfs2002q3	25/1/2561 10:03	Stata Dataset	25 KB
 lfs2002q4	25/1/2561 10:03	Stata Dataset	2,131 KB
 lfs2003q1	25/1/2561 10:03	Stata Dataset	2,130 KB
 lfs2003q2	25/1/2561 10:03	Stata Dataset	18 KB
 lfs2003q3	25/1/2561 10:03	Stata Dataset	17 KB
 lfs2003q4	25/1/2561 10:03	Stata Dataset	2,106 KB

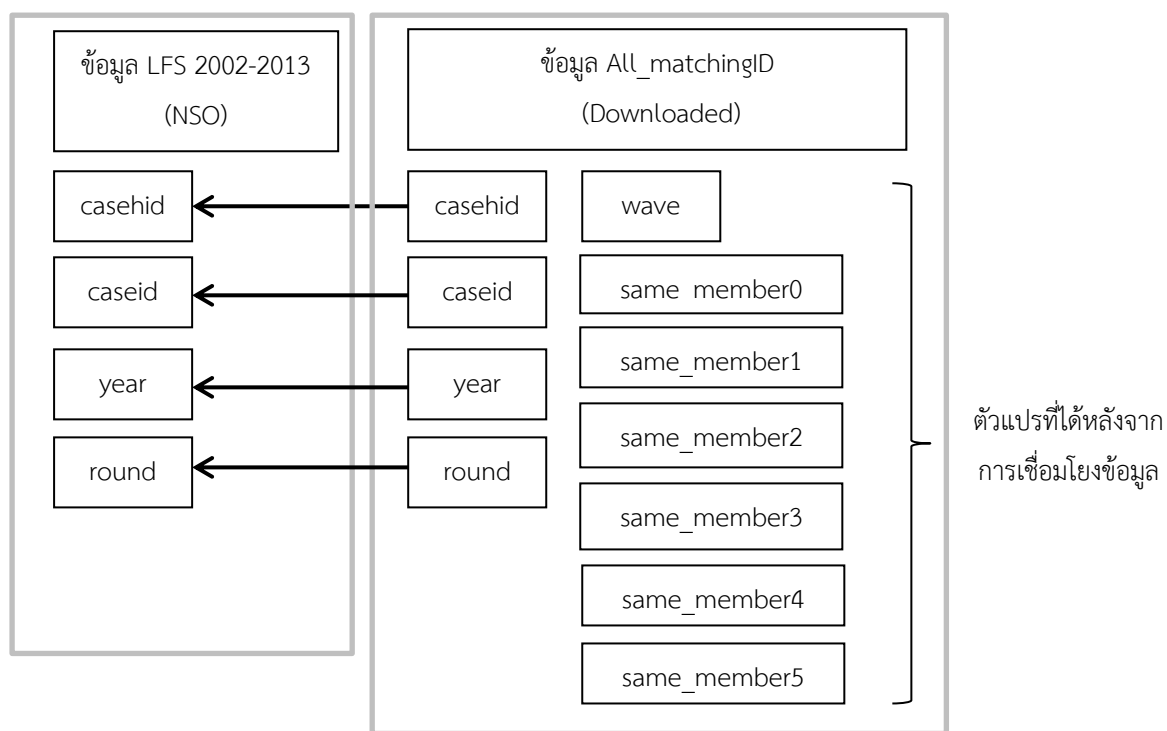
3. ดาวน์โหลดข้อมูล matching ID ที่ชื่อ All_matchingID ใส่ไว้ในโฟลเดอร์ชื่อ data และ ดาวน์โหลด command (Stata dofile) ใส่ไว้ในโฟลเดอร์ชื่อ dofile ผ่านเว็บไซต์ที่เผยแพร่ โดย command ที่โหลดมาพร้อมกับข้อมูล All_matchingID จะใช้สร้างตัวแปร ID เพื่อเชื่อมโยงข้อมูลภาวะการทำงานของประชากร

ข้อมูล matching ID แยกรายชุด

Name	Date modified	Type	Size
 All_matchingID_2002-2003_ORG	17/1/2561 15:32	Stata Dataset	220,702 KB
 All_matchingID_2004-2006_ORG	17/1/2561 15:33	Stata Dataset	293,803 KB
 All_matchingID_2006-2008_ORG	17/1/2561 15:33	Stata Dataset	352,003 KB
 All_matchingID_2008-2010_ORG	17/1/2561 15:33	Stata Dataset	344,633 KB
 All_matchingID_2010-2011_ORG	17/1/2561 15:33	Stata Dataset	344,633 KB
 All_matchingID_2012-2013_ORG	17/1/2561 15:33	Stata Dataset	238,147 KB

Stata dofile แยกรายชุด

 do-matchingID_2002-2003	25/1/2561 9:59	Stata Do-file	4 KB
 do-matchingID_2004-2006	25/1/2561 12:24	Stata Do-file	4 KB
 do-matchingID_2006-2008	25/1/2561 13:21	Stata Do-file	4 KB
 do-matchingID_2008-2010	25/1/2561 13:26	Stata Do-file	4 KB
 do-matchingID_2010-2011	25/1/2561 13:28	Stata Do-file	4 KB
 do-matchingID_2012-2013	25/1/2561 13:29	Stata Do-file	4 KB



แผนภาพการเชื่อมโยงโดยใช้ Matching ID

4. เปิดข้อมูลภาวะการทำงานของประชากร แล้วใช้ Stata dofile ที่ดาวน์โหลดมาในขั้นตอนที่ 2 สร้างตัวแปร ID ครัวเรือน และ ID รายบุคคล เพื่อให้สามารถเชื่อมโยง ID ดังกล่าวเข้ากับ ข้อมูล All_matchingID_2002-2013_ORG โดยใช้ตัวแปร casehid (ID ครัวเรือน), caseid (ID รายบุคคล), year (ปี) และ round (ไตรมาส) ในการเชื่อมโยง ซึ่งตัวแปรที่ใช้สร้าง ID ครัวเรือน และ ID รายบุคคล ประกอบด้วยตัวแปร ดังนี้

- ID ครัวเรือน ให้สร้างแล้วตั้งชื่อว่า casehid

casehid = region + province + areatype + psunumbr + samset + ea_set +
enumdist + blk_vil + hhtype + hhnumber

ตัวอย่าง casehid = 11010001A1060000101

ตัวแปร	คำอธิบาย	ตัวอย่าง	หลักที่
region	ภาค	1	1
province	จังหวัด	10	2-3
areatype	อยู่ในเขตเทศบาล/นอกเขตเทศบาล	1	4
psunumbr	ลำดับที่ชุมนุมอาคาร/หมู่บ้านตัวอย่าง	0001	5-8
samset	ชุดครัวเรือนตัวอย่าง	A	9
ea_set	ลำดับที่ของชุดครัวเรือนตัวอย่าง	1	10
enumdist	เขตที่แน่นับ	060	11-13
blk_vil	หมู่บ้าน	000	14-16
hhtype	ประเภทของครัวเรือน	1	17
hhnumber	หมายเลขครัวเรือน	01	18-19

- ID รายบุคคล ให้สร้างแล้วตั้งชื่อว่า caseid

casehid = region + province + areatype + psunumbr + samset + ea_set +
enumdist + blk_vil + hhtype + hhnumber + memnumbr

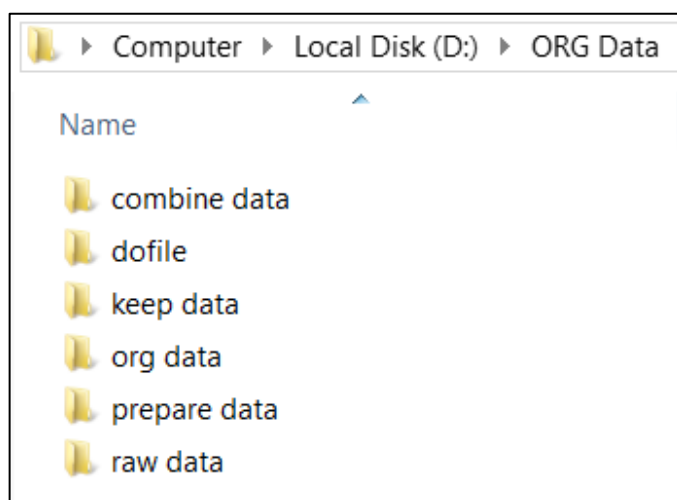
ตัวอย่าง caseid = 11010001A106000010101

ตัวแปร	คำอธิบาย	ตัวอย่าง	หลักที่
region	ภาค	1	1
province	จังหวัด	10	2-3
areatype	อยู่ในเขตเทศบาล/นอกเขตเทศบาล	1	4
psunumbr	ลำดับที่ชุมนุมอาคาร/หมู่บ้านตัวอย่าง	0001	5-8
samset	ชุดครัวเรือนตัวอย่าง	A	9
ea_set	ลำดับที่ของชุดครัวเรือนตัวอย่าง	1	10
enumdist	เขตที่แน่นับ	060	11-13
blk_vil	หมู่บ้าน	000	14-16
hhtype	ประเภทของครัวเรือน	1	17
hhnumber	หมายเลขครัวเรือน	01	18-19
memnumbr	หมายเลขสมาชิกครัวเรือน	01	20-21

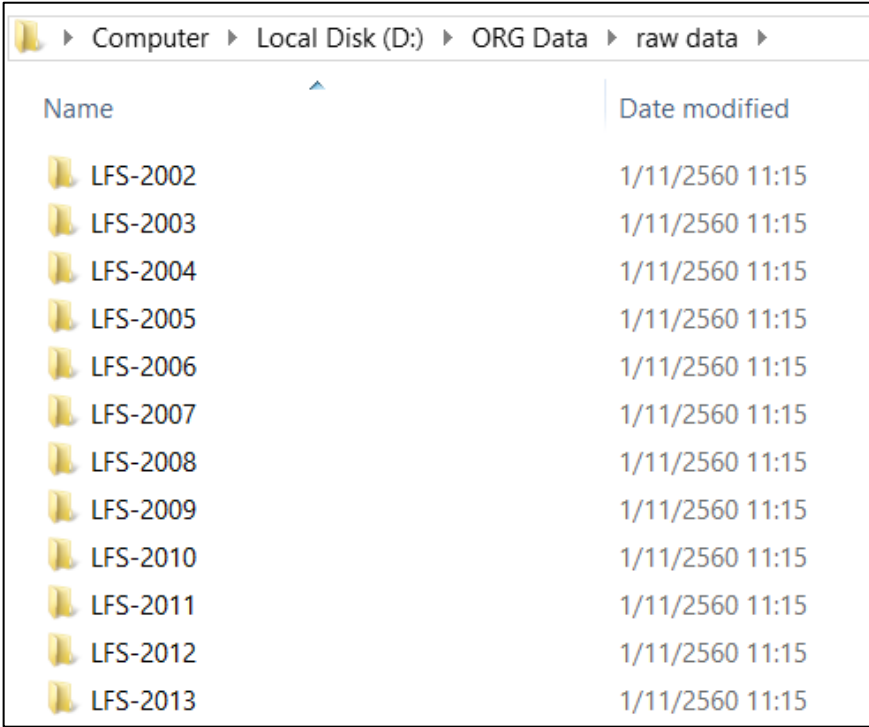
5. หลังจากเชื่อมโยงข้อมูลเข้าด้วยกัน ก็จะได้ข้อมูลตัวอย่างซ้ำจากข้อมูลภาวะการทำงานของประชากรโดยมีตัวแปรที่ระบุว่า สมาชิกครัวเรือนคนใดที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง ซึ่งประกอบไปด้วยตัวแปรที่ชื่อ same_member0, same_member1, same_member2, same_member3, same_member4 และ same_member5 ตามลำดับ โดยตั้งชื่อหมายเลขตามเกณฑ์ที่นักวิจัยใช้ในการค้นหาข้อมูลตัวอย่างซ้ำรายบุคคล ซึ่งจะอธิบายรายละเอียดของแต่ละเกณฑ์ในส่วนถัดไป

3. วิธีที่ 2 ขั้นตอนการใช้โค้ดคำสั่ง เพื่อสร้างข้อมูล ORG

1. เตรียมข้อมูลภาวะการทำงานของประชากรจากสำนักงานสถิติแห่งชาติ ตั้งแต่ปี 2002-2013 (พ.ศ. 2545-2556) ซึ่งมีทั้งไฟล์ในรูปแบบ .sav (SPSS file) และ .dat (Text file)
2. ดาวน์โหลดโค้ดคำสั่งที่อยู่ในรูปแบบไฟล์ .do ซึ่งต้องเปิดกับโปรแกรม Stata
3. เตรียมโฟลเดอร์เพื่อจัดเก็บข้อมูล ซึ่งประกอบไปด้วยโฟลเดอร์ต่างๆ ดังต่อไปนี้
สร้างโฟลเดอร์หลักตั้งชื่อว่า ORG Data ภายในโฟลเดอร์หลักจะประกอบไปด้วย 6 โฟลเดอร์ย่อย คือ
 - 1) raw data - ใส่ข้อมูล .dat ที่ได้จาก NSO แยกรายปี
 - 2) prepare data
 - 3) combine data
 - 4) keep data
 - 5) org data
 - 6) dofile - ใส่ dofile ที่ load มา

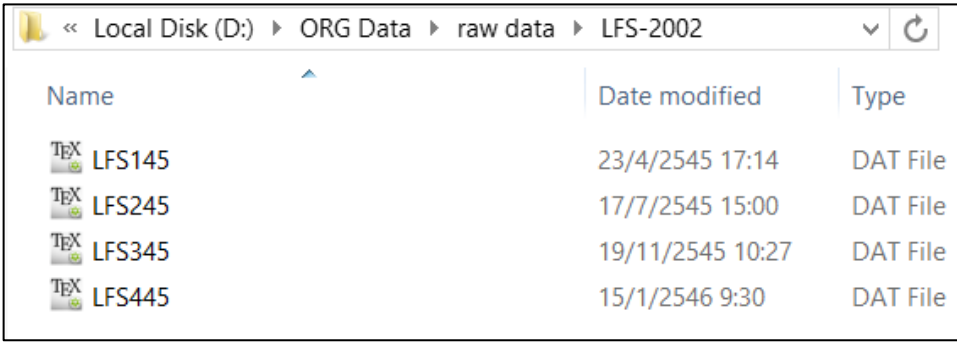


4. ใน folder raw data ให้ใส่ข้อมูลที่เป็นไฟล์ .dat โดยตั้งชื่อไฟล์เดอร์และชื่อ data ดังนี้



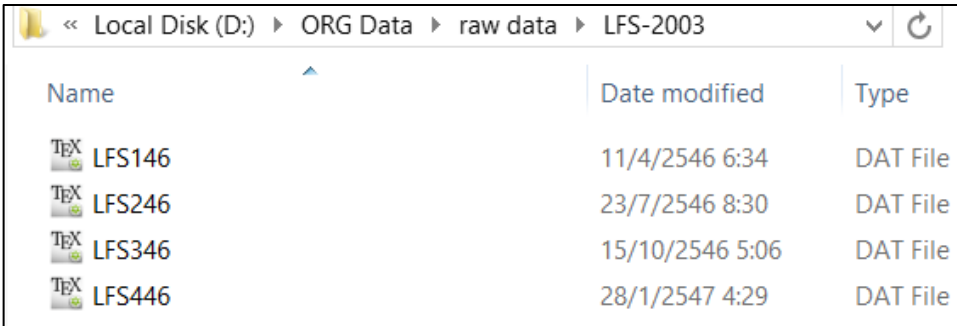
Name	Date modified
LFS-2002	1/11/2560 11:15
LFS-2003	1/11/2560 11:15
LFS-2004	1/11/2560 11:15
LFS-2005	1/11/2560 11:15
LFS-2006	1/11/2560 11:15
LFS-2007	1/11/2560 11:15
LFS-2008	1/11/2560 11:15
LFS-2009	1/11/2560 11:15
LFS-2010	1/11/2560 11:15
LFS-2011	1/11/2560 11:15
LFS-2012	1/11/2560 11:15
LFS-2013	1/11/2560 11:15

- โฟล์เดอร์ LFS-2002 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้



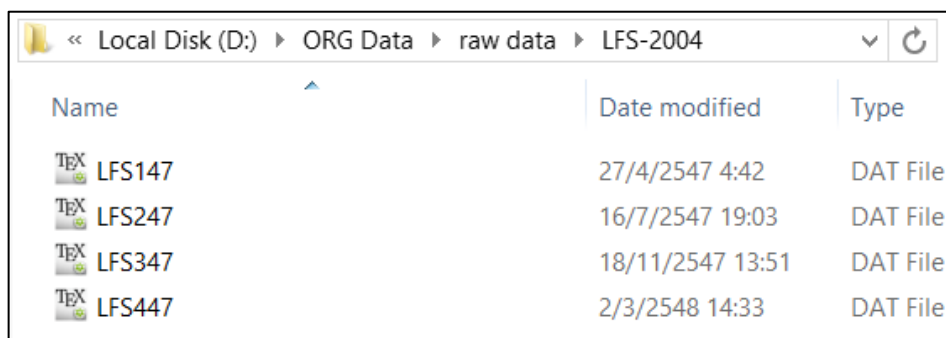
Name	Date modified	Type
LFS145	23/4/2545 17:14	DAT File
LFS245	17/7/2545 15:00	DAT File
LFS345	19/11/2545 10:27	DAT File
LFS445	15/1/2546 9:30	DAT File

- โฟล์เดอร์ LFS-2003 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้



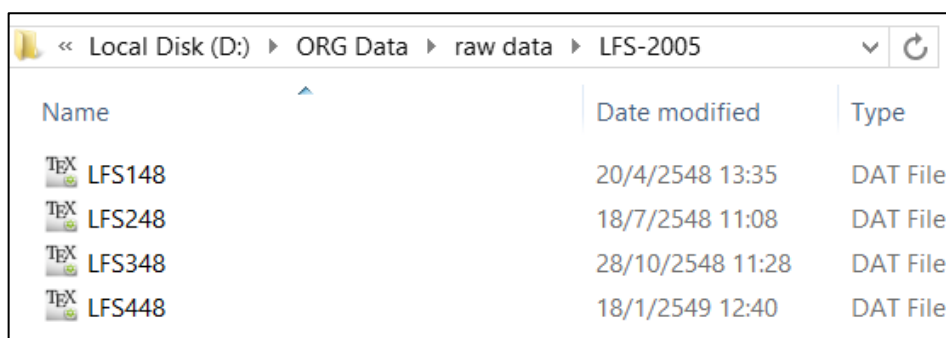
Name	Date modified	Type
LFS146	11/4/2546 6:34	DAT File
LFS246	23/7/2546 8:30	DAT File
LFS346	15/10/2546 5:06	DAT File
LFS446	28/1/2547 4:29	DAT File

- โพล์เตอร์ LFS-2004 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้



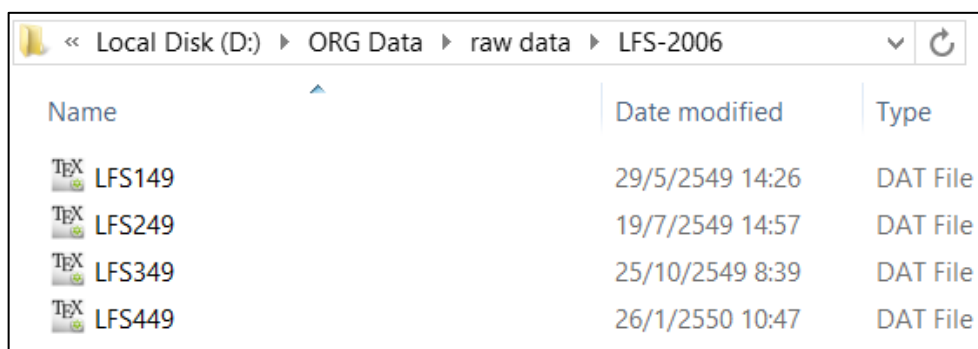
Name	Date modified	Type
 LFS147	27/4/2547 4:42	DAT File
 LFS247	16/7/2547 19:03	DAT File
 LFS347	18/11/2547 13:51	DAT File
 LFS447	2/3/2548 14:33	DAT File

- โพล์เตอร์ LFS-2005 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้



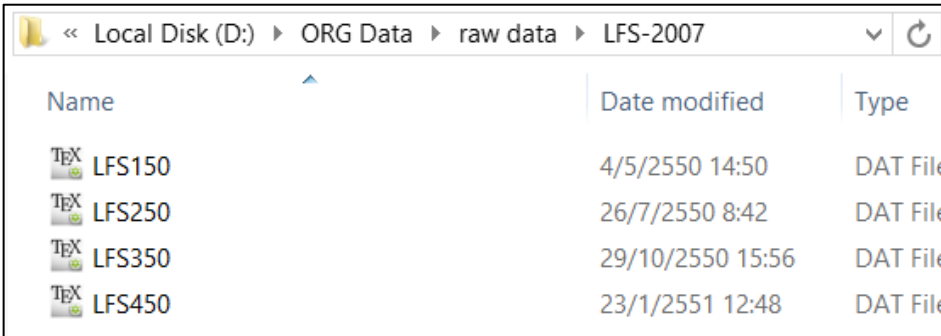
Name	Date modified	Type
 LFS148	20/4/2548 13:35	DAT File
 LFS248	18/7/2548 11:08	DAT File
 LFS348	28/10/2548 11:28	DAT File
 LFS448	18/1/2549 12:40	DAT File

- โพล์เตอร์ LFS-2006 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้



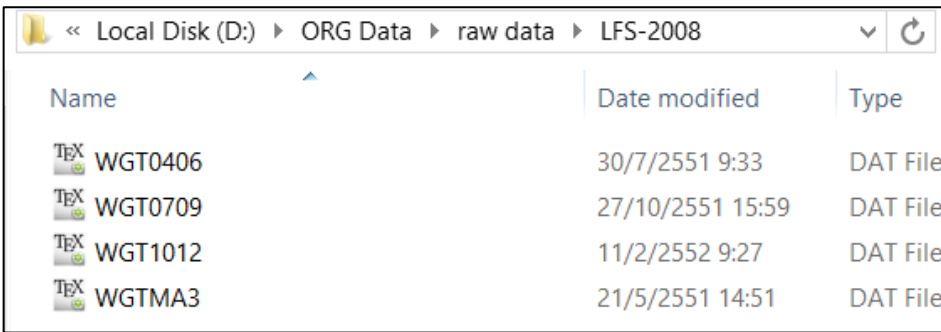
Name	Date modified	Type
 LFS149	29/5/2549 14:26	DAT File
 LFS249	19/7/2549 14:57	DAT File
 LFS349	25/10/2549 8:39	DAT File
 LFS449	26/1/2550 10:47	DAT File

- โพล์เตอร์ LFS-2007 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้



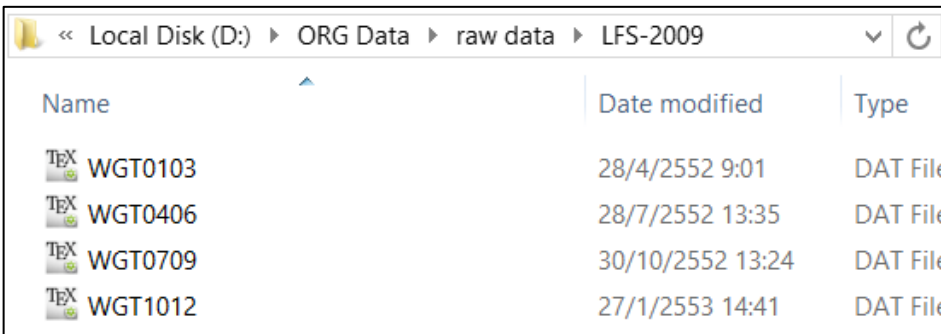
Name	Date modified	Type
 LFS150	4/5/2550 14:50	DAT File
 LFS250	26/7/2550 8:42	DAT File
 LFS350	29/10/2550 15:56	DAT File
 LFS450	23/1/2551 12:48	DAT File

- โพล์เตอร์ LFS-2008 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้



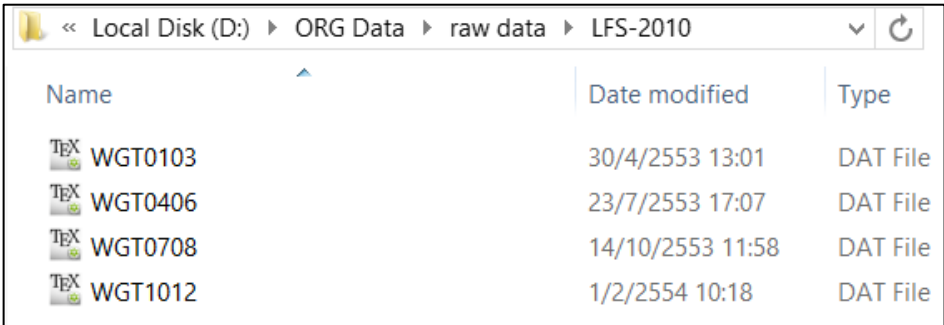
Name	Date modified	Type
 WGT0406	30/7/2551 9:33	DAT File
 WGT0709	27/10/2551 15:59	DAT File
 WGT1012	11/2/2552 9:27	DAT File
 WGTMA3	21/5/2551 14:51	DAT File

- โพล์เตอร์ LFS-2009 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้



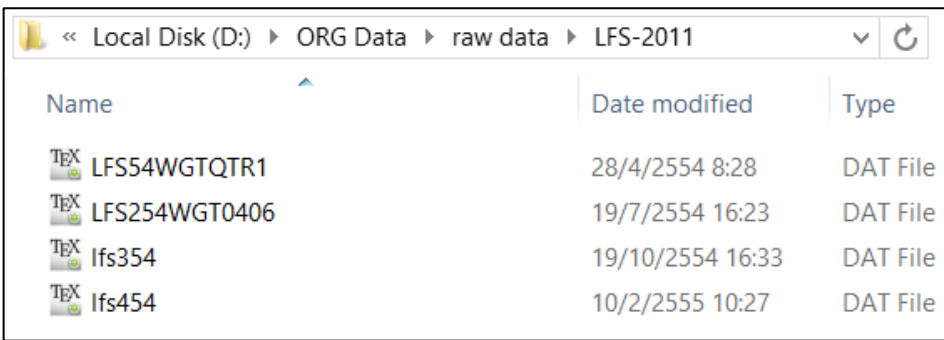
Name	Date modified	Type
 WGT0103	28/4/2552 9:01	DAT File
 WGT0406	28/7/2552 13:35	DAT File
 WGT0709	30/10/2552 13:24	DAT File
 WGT1012	27/1/2553 14:41	DAT File

- โพล์เดอร์ LFS-2010 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้



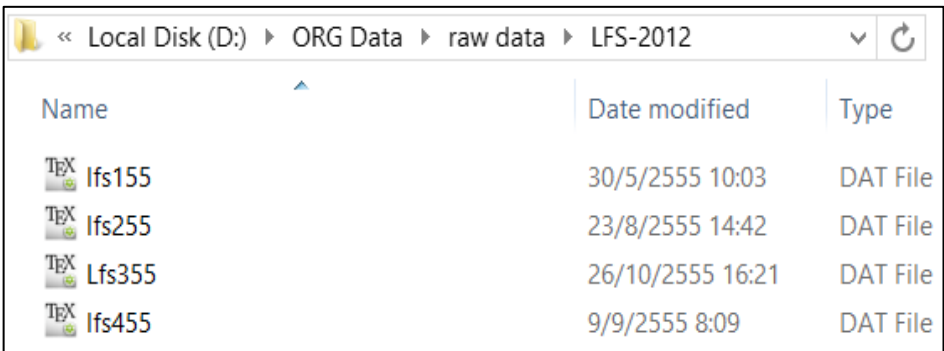
Name	Date modified	Type
WGT0103	30/4/2553 13:01	DAT File
WGT0406	23/7/2553 17:07	DAT File
WGT0708	14/10/2553 11:58	DAT File
WGT1012	1/2/2554 10:18	DAT File

- โพล์เดอร์ LFS-2011 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้



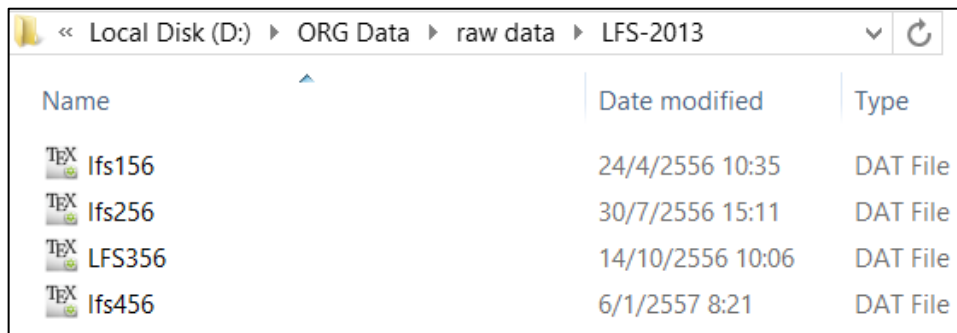
Name	Date modified	Type
LFS54WGTQTR1	28/4/2554 8:28	DAT File
LFS254WGT0406	19/7/2554 16:23	DAT File
lfs354	19/10/2554 16:33	DAT File
lfs454	10/2/2555 10:27	DAT File

- โพล์เดอร์ LFS-2012 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้



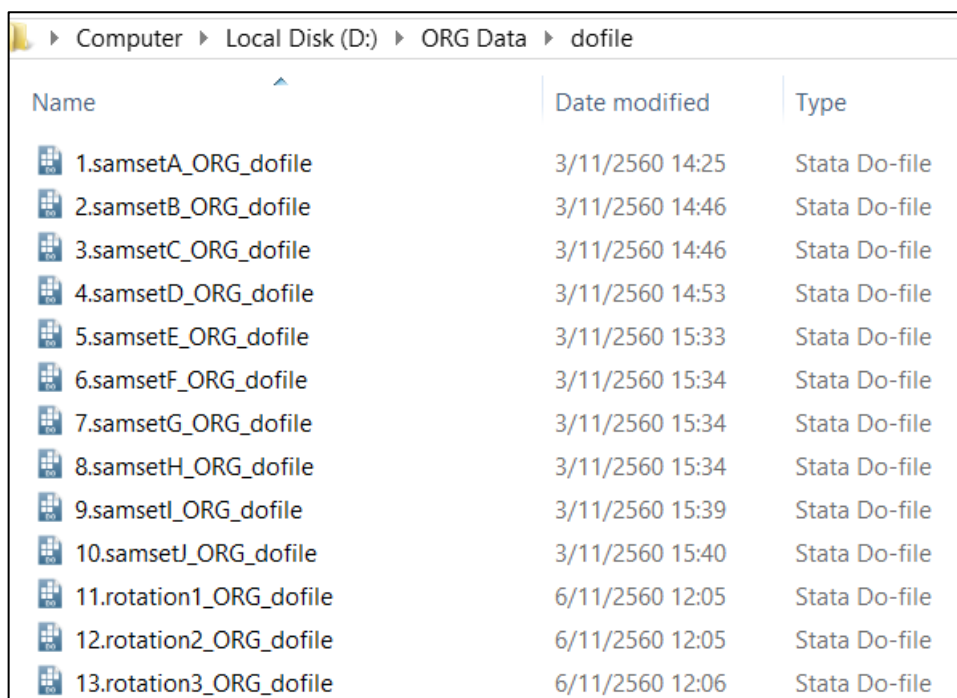
Name	Date modified	Type
lfs155	30/5/2555 10:03	DAT File
lfs255	23/8/2555 14:42	DAT File
Lfs355	26/10/2555 16:21	DAT File
lfs455	9/9/2555 8:09	DAT File

- โพลเดอร์ LFS-2013 ใส่ raw data ที่อยู่ในรูปแบบไฟล์ .dat โดยให้ตั้งชื่อเรียงตามไตรมาสที่ 1-4 ดังนี้



Name	Date modified	Type
lfs156	24/4/2556 10:35	DAT File
lfs256	30/7/2556 15:11	DAT File
LFS356	14/10/2556 10:06	DAT File
lfs456	6/1/2557 8:21	DAT File

5. ดาวน์โหลด dofile ทั้งหมด 13 ไฟล์ ไปใส่ไว้ในโพลเดอร์ dofile

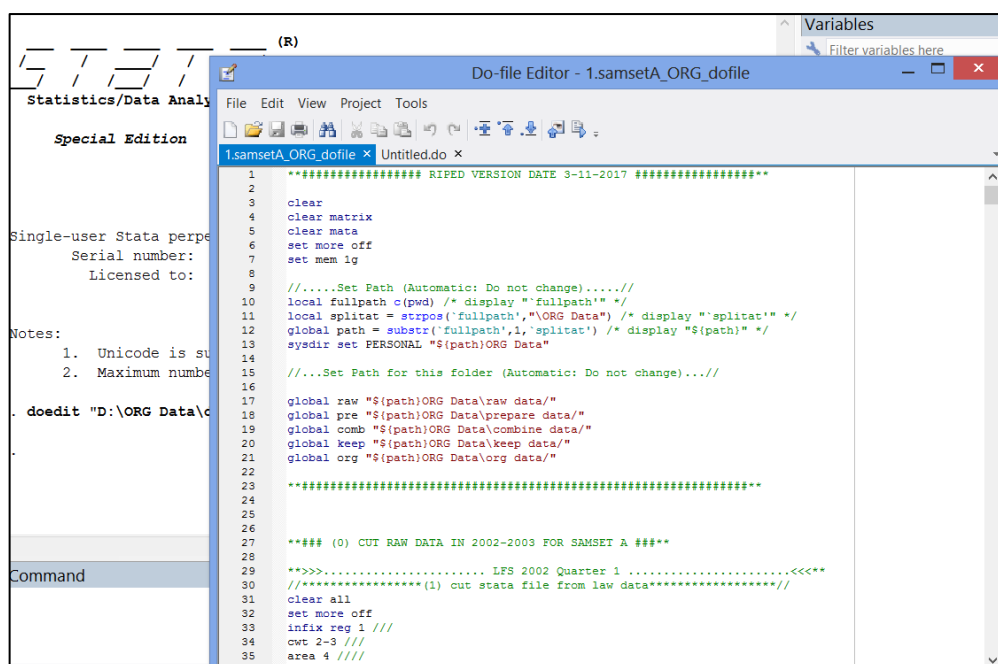


Name	Date modified	Type
1.samsetA_ORG_dofile	3/11/2560 14:25	Stata Do-file
2.samsetB_ORG_dofile	3/11/2560 14:46	Stata Do-file
3.samsetC_ORG_dofile	3/11/2560 14:46	Stata Do-file
4.samsetD_ORG_dofile	3/11/2560 14:53	Stata Do-file
5.samsetE_ORG_dofile	3/11/2560 15:33	Stata Do-file
6.samsetF_ORG_dofile	3/11/2560 15:34	Stata Do-file
7.samsetG_ORG_dofile	3/11/2560 15:34	Stata Do-file
8.samsetH_ORG_dofile	3/11/2560 15:34	Stata Do-file
9.samsetI_ORG_dofile	3/11/2560 15:39	Stata Do-file
10.samsetJ_ORG_dofile	3/11/2560 15:40	Stata Do-file
11.rotation1_ORG_dofile	6/11/2560 12:05	Stata Do-file
12.rotation2_ORG_dofile	6/11/2560 12:05	Stata Do-file
13.rotation3_ORG_dofile	6/11/2560 12:06	Stata Do-file

- dofile ที่ใช้ข้อมูล LFS ในปี 2002-2003 คือ
 - 1.samsetA_ORG_dofile.do
 - 2.samsetB_ORG_dofile.do
- dofile ที่ใช้ข้อมูล LFS ในปี 2004-2006 คือ
 - 3.samsetC_ORG_dofile.do
 - 4.samsetD_ORG_dofile.do

- dofile ที่ใช้ข้อมูล LFS ในปี 2006-2008 คือ
 - 5.samsetE_ORG_dofile.do
 - 6.samsetF_ORG_dofile.do
- dofile ที่ใช้ข้อมูล LFS ในปี 2008-2010 คือ
 - 7.samsetG_ORG_dofile.do
 - 8.samsetH_ORG_dofile.do
- dofile ที่ใช้ข้อมูล LFS ในปี 2010-2011 คือ
 - 9.samsetI_ORG_dofile.do
 - 10.samsetJ_ORG_dofile.do
- dofile ที่ใช้ข้อมูล LFS ในปี 2012-2013 คือ
 - 11.rotation1_ORG_dofile.do
 - 12.rotation2_ORG_dofile.do
 - 13.rotation3_ORG_dofile.do

6. เปิด dofile แต่ละชุด รันตามหมายเลข 1-13



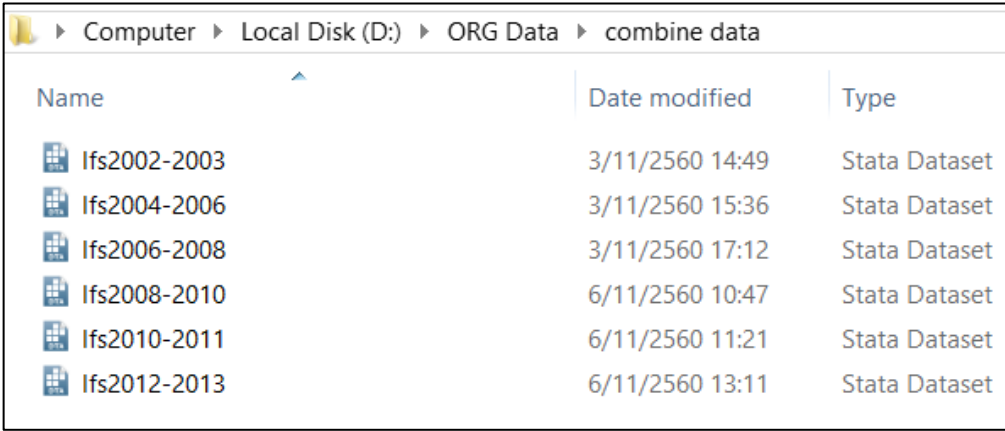
7. ข้อมูล LFS ที่ยังไม่มีจัดการให้เป็นข้อมูลตัวอย่างซ้ำ ORG แต่จะเป็นข้อมูลที่มีการสร้างชื่อตัวแปรให้เหมือนกันในทุกปีและทุกไตรมาสจะถูกเก็บอยู่ในโฟลเดอร์ prepare data

Name	Date modified	Type
lfs2002q1	3/11/2560 14:48	Stata Dataset
lfs2002q2	3/11/2560 14:48	Stata Dataset
lfs2002q3	3/11/2560 14:48	Stata Dataset
lfs2002q4	3/11/2560 14:48	Stata Dataset
lfs2003q1	3/11/2560 14:49	Stata Dataset
lfs2003q2	3/11/2560 14:49	Stata Dataset
lfs2003q3	3/11/2560 14:49	Stata Dataset
lfs2003q4	3/11/2560 14:49	Stata Dataset
lfs2004q1	3/11/2560 15:33	Stata Dataset
lfs2004q2	3/11/2560 15:33	Stata Dataset
lfs2004q3	3/11/2560 15:33	Stata Dataset
lfs2004q4	3/11/2560 15:34	Stata Dataset

8. ข้อมูล LFS ที่แบ่งตาม samset และ rotation แต่ยังไม่มีการจัดการให้เป็นข้อมูล ORG ตัวอย่างซ้ำ จะถูกเก็บไว้ในโฟลเดอร์ keep data

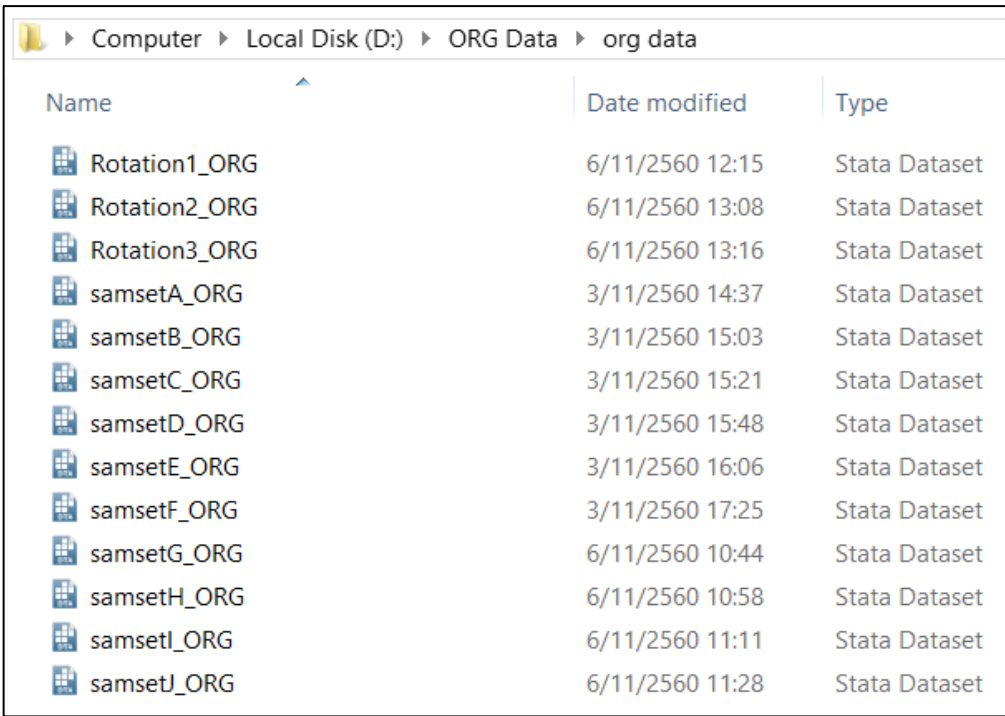
Name	Date modified	Type
Rotation1_samsetA	6/11/2560 12:08	Stata Dataset
Rotation1_samsetA_keep	6/11/2560 12:11	Stata Dataset
Rotation1_samsetB	6/11/2560 12:10	Stata Dataset
Rotation1_samsetB_keep	6/11/2560 12:11	Stata Dataset
Rotation1_samsetC	6/11/2560 12:10	Stata Dataset
Rotation1_samsetC_keep	6/11/2560 12:11	Stata Dataset
Rotation1_samsetD	6/11/2560 12:11	Stata Dataset
Rotation1_samsetD_keep	6/11/2560 12:12	Stata Dataset

9. ข้อมูล LFS ที่รวมชุดข้อมูลในแต่ละช่วงก่อนที่จะนำไปจัดการให้เป็นข้อมูล ORG จะเก็บอยู่ในโฟลเดอร์ combine data



Name	Date modified	Type
lfs2002-2003	3/11/2560 14:49	Stata Dataset
lfs2004-2006	3/11/2560 15:36	Stata Dataset
lfs2006-2008	3/11/2560 17:12	Stata Dataset
lfs2008-2010	6/11/2560 10:47	Stata Dataset
lfs2010-2011	6/11/2560 11:21	Stata Dataset
lfs2012-2013	6/11/2560 13:11	Stata Dataset

10. ข้อมูล ORG แต่ละชุดจะอยู่ในโฟลเดอร์ org data



Name	Date modified	Type
Rotation1_ORG	6/11/2560 12:15	Stata Dataset
Rotation2_ORG	6/11/2560 13:08	Stata Dataset
Rotation3_ORG	6/11/2560 13:16	Stata Dataset
samsetA_ORG	3/11/2560 14:37	Stata Dataset
samsetB_ORG	3/11/2560 15:03	Stata Dataset
samsetC_ORG	3/11/2560 15:21	Stata Dataset
samsetD_ORG	3/11/2560 15:48	Stata Dataset
samsetE_ORG	3/11/2560 16:06	Stata Dataset
samsetF_ORG	3/11/2560 17:25	Stata Dataset
samsetG_ORG	6/11/2560 10:44	Stata Dataset
samsetH_ORG	6/11/2560 10:58	Stata Dataset
samsetI_ORG	6/11/2560 11:11	Stata Dataset
samsetJ_ORG	6/11/2560 11:28	Stata Dataset

ตารางที่ 1: ตัวอย่างตารางข้อมูล ORG สำหรับ samset A

caseid	caseid_new	caseid_new2	wave	sex	age	same_member0	same_member1	same_member2	same_member3	same_member4
11010001A100000010501	11010001A100000010501	11010001A100000010501	3	2	33	same	same	same	same	same
11010001A100000010501	11010001A100000010501	11010001A100000010501	4	2	33	same	same	same	same	same
11010001A100000010502	11010001A100000010502	11010001A100000010502	1	2	21	same	same	same	same	not same
11010001A100000010502	11010001A100000010502	11010001A100000010502	2	2	21	same	same	same	same	not same
11010001A100000010502	11010001A100000010502	11010001A100000010502	3	2	22	same	same	same	same	not same
11010001A100000010502	11010001A100000010502	11010001A100000010502	4	2	22	same	same	same	same	not same
11010001A100000010601	11010001A100000010606	11010001A100000010606	2	1	49	not same	not same	not same	not same	not same
11010001A100000010601	11010001A100000010609	11010001A100000010609	4	1	46	not same	not same	not same	not same	not same
11010001A100000010602	11010001A100000010607	11010001A100000010607	2	2	25	not same	same	same	same	same
11010001A100000010602	11010001A100000010610	11010001A100000010610	4	2	32	not same	not same	not same	not same	not same
11010001A100000010603	11010001A100000010608	11010001A100000010608	2	1	20	not same	same	same	same	not same
11010001A100000010603	11010001A100000010607	11010001A100000010607	4	2	26	not same	same	same	same	not same
11010001A100000010604	11010001A100000010608	11010001A100000010608	4	1	21	not same	same	same	same	same
11010001A100000010605	11010001A100000010611	11010001A100000010611	4	2	1	not same	not same	not same	not same	not same
11010001A100000010701	11010001A100000010701	11010001A100000010701	1	1	54	same	same	same	same	same
11010001A100000010701	11010001A100000010701	11010001A100000010701	2	1	54	same	same	same	same	same
11010001A100000010701	11010001A100000010701	11010001A100000010701	3	1	55	same	same	same	same	same
11010001A100000010701	11010001A100000010701	11010001A100000010701	4	1	56	same	same	not same	not same	not same
11010001A100000010702	11010001A100000010702	11010001A100000010702	1	2	44	same	same	same	same	same
11010001A100000010702	11010001A100000010702	11010001A100000010702	2	2	44	same	same	same	same	same
11010001A100000010702	11010001A100000010702	11010001A100000010702	3	2	45	same	same	same	same	same
11010001A100000010702	11010001A100000010702	11010001A100000010702	4	2	45	same	same	same	same	same
11010001A100000010703	11010001A100000010703	11010001A100000010703	1	2	26	same	same	same	same	same
11010001A100000010703	11010001A100000010703	11010001A100000010703	2	2	26	same	same	same	same	same
11010001A100000010703	11010001A100000010703	11010001A100000010703	3	2	27	same	same	same	same	same

4. เกณฑ์ที่ใช้ในการค้นหาข้อมูลตัวอย่างซ้ำรายบุคคล แบ่งออกเป็น 5 เกณฑ์ ได้แก่

6. กลุ่มตัวอย่างซ้ำกันต้องมีเพศเดียวกัน และอายุไม่เกิน 2 ปี นับตั้งแต่การถูกสัมภาษณ์ ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่ สัมภาษณ์ (ตัวแปรที่ใช้ same_member1)
7. กลุ่มตัวอย่างซ้ำกันต้องมีเพศเดียวกัน และอายุไม่เกิน 1 ปี นับตั้งแต่การถูกสัมภาษณ์ ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่ สัมภาษณ์ (ตัวแปรที่ใช้ same_member2)
8. กลุ่มตัวอย่างซ้ำกันต้องมีเพศเดียวกัน และอายุไม่เกิน 1 ปี นับตั้งแต่การถูกสัมภาษณ์ ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่ สัมภาษณ์ และสถานภาพสมรสต้องเป็นไปตามเงื่อนไขของแต่ละสถานภาพ (ตัวแปรที่ใช้ same_member3)
9. กลุ่มตัวอย่างซ้ำกันต้องมีเพศเดียวกัน และอายุไม่เกิน 1 ปี นับตั้งแต่การถูกสัมภาษณ์ ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่ สัมภาษณ์ และจำนวนปีการศึกษาจะต้องมีจำนวนเท่าเดิมหรือเพิ่มขึ้นไม่เกิน 1 ปีการศึกษา นับตั้งแต่การถูกสัมภาษณ์ครั้งแรกจนถึงครั้งสุดท้าย (ตัวแปรที่ใช้ same_member4)
10. กลุ่มตัวอย่างซ้ำกันต้องมีเพศเดียวกัน และอายุไม่เกิน 1 ปี นับตั้งแต่การถูกสัมภาษณ์ ครั้งแรกไปจนถึงครั้งสุดท้ายที่ถูกสัมภาษณ์ และอายุจะต้องเท่าเดิมหรือเพิ่มขึ้นตามครั้งที่ สัมภาษณ์ และสถานภาพสมรสเป็นไปตามเงื่อนไข และจำนวนปีการศึกษาจะต้องมีจำนวน เท่าเดิมหรือเพิ่มขึ้นไม่เกิน 1 ปีการศึกษานับตั้งแต่การถูกสัมภาษณ์ครั้งแรกจนถึงครั้งสุดท้าย (ตัวแปรที่ใช้ same_member5)

หมายเหตุ : เกณฑ์ที่ 0 เป็นเกณฑ์ที่ใช้แยกระหว่างการเชื่อมโยงโดยใช้ caseid และไม่สามารถใช้ caseid ได้ ซึ่งต้องเชื่อมโยงจากการใช้เพศเดียวกันและอายุที่ใกล้เคียงกัน (ตัวแปรที่ใช้ same_member0)

นอกจากนี้ นักวิจัยได้ทำการตรวจสอบข้อมูลตัวอย่างซ้ำรายบุคคลโดยใช้ caseid ที่สร้างขึ้นจากตัวแปร demographics และ caseid ใหม่ที่สร้างขึ้นจากข้อมูลที่ไม่สามารถใช้ caseid เดิมได้ โดยตั้งชื่อว่า caseid_new เพราะพบว่า อายุและเพศต่างจาก caseid ที่ถูกสัมภาษณ์ในครั้งแรก ตามตัวอย่างในตารางที่ 2

ตารางที่ 2: ตัวอย่างตารางข้อมูลที่ต้องมีการสร้าง caseid ใหม่

ครัวเรือนที่	ลำดับที่สมาชิก	ปี	ไตรมาส	ครั้งที่สัมภาษณ์	เพศ	อายุ	caseid
1	1	2545	1	1	ชาย	55	ใช้แบบเดิม
1	1	2545	4	2	ชาย	34	ต้องสร้างใหม่
1	1	2546	1	3	ชาย	34	ต้องสร้างใหม่
1	1	2546	4	4	ชาย	35	ต้องสร้างใหม่
1	2	2545	1	1	หญิง	29	ใช้แบบเดิม
1	2	2545	4	2	หญิง	29	ใช้แบบเดิม
1	2	2546	1	3	ชาย	17	ต้องสร้างใหม่
1	2	2546	4	4	ชาย	18	ต้องสร้างใหม่

จากตารางที่ 2 จะเห็นได้ว่า ครัวเรือนตัวอย่างที่ 1 สมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 1 มีอายุ 55 ปี เพศชาย และเมื่อสมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 2 มีอายุ 34 ปี เพศชาย และเมื่อถูกสัมภาษณ์ครั้งที่ 3 และ 4 สมาชิกคนที่ 1 มีอายุ 34 ปี เพศชาย และอายุ 35 ปี เพศชาย ตามลำดับ ซึ่งพบว่า สมาชิกคนที่ 1 ที่ถูกสัมภาษณ์ครั้งแรกมีอายุห่างจากการถูกสัมภาษณ์ครั้งที่ 2-4 มากพอสมควร ทำให้สามารถสรุปได้ว่า สมาชิกคนที่ 1 ที่ถูกสัมภาษณ์ครั้งที่ 1 ไม่ใช่คนเดียวกับสมาชิกคนที่ 1 ที่ถูกสัมภาษณ์ครั้งที่ 2-4 เพราะฉะนั้น จึงจำเป็นต้องสร้าง caseid ใหม่ให้กับสมาชิกที่ถูกสัมภาษณ์ครั้งที่ 2-4

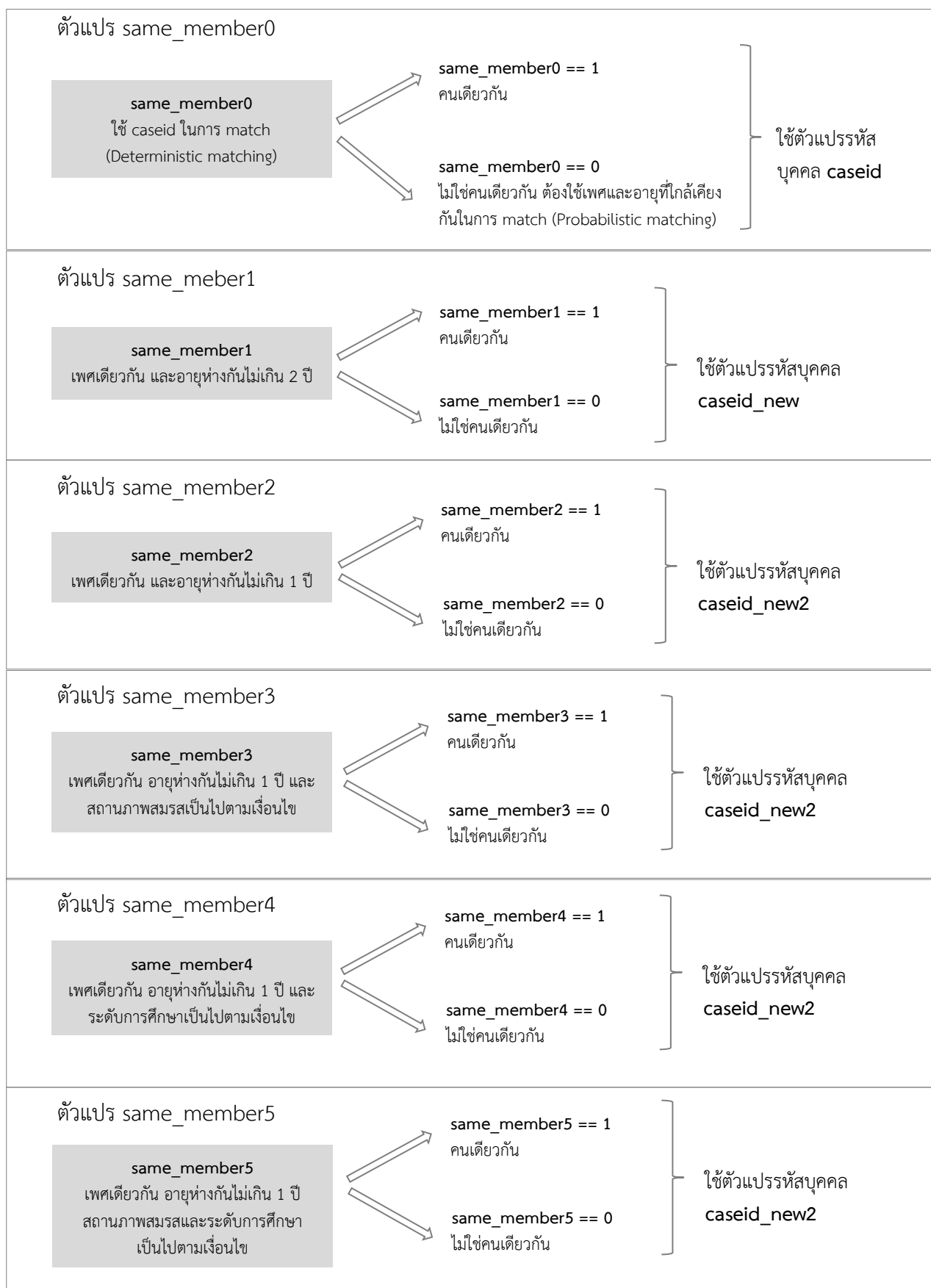
ในการตรวจสอบข้อมูลกลุ่มตัวอย่างซ้ำที่ใช้ caseid ใหม่ ก็ยังคงใช้เกณฑ์ตามเงื่อนไขที่ 1 – 5 เช่นเดียวกับการใช้ caseid แบบเดิม ซึ่งสำหรับการสร้าง caseid ใหม่ นักวิจัยใช้ตัวแปรเพศและอายุในการเชื่อมโยง โดยได้แบ่งการสร้าง caseid ใหม่จาก 2 เงื่อนไข คือ 1) อายุของคนดังกล่าวที่ต้องใช้ caseid_new เดียวกันจะต้องมีอายุห่างกันไม่เกิน 2 ปี (กำหนดเป็น caseid_new) และ 2) อายุของคนดังกล่าวที่ต้องใช้ caseid_new เดียวกันจะต้องมีอายุห่างกันไม่เกิน 1 ปี (กำหนดเป็น caseid_new2) นอกจากนี้ นักวิจัยจะใช้เลข 2 หลักสุดท้ายของ caseid_new ซึ่งเป็นลำดับที่ต่อจากสมาชิกครัวเรือนสุดท้ายของ caseid เดิม ตามตัวอย่างในตารางที่ 3 ดังต่อไปนี้

ตารางที่ 3: ตัวอย่างตารางข้อมูลที่มีการสร้าง caseid_new และ caseid_new2

ครัวเรือนที่	ลำดับที่สมาชิก	ปี	ไตรมาส	ครั้งที่สัมภาษณ์	เพศ	อายุ	caseid	caseid_new	caseid_new2
2	1	2545	1	1	ชาย	42	01	01	01
2	1	2545	4	2	ชาย	42	01	01	01
2	1	2546	1	3	ชาย	35	01	03	03
2	1	2546	4	4	ชาย	37	01	03	04
2	2	2545	1	1	หญิง	29	02	02	02
2	2	2545	4	2	หญิง	29	02	02	02
2	2	2546	1	3	ชาย	18	02	04	05
2	2	2546	4	4	ชาย	18	02	04	05

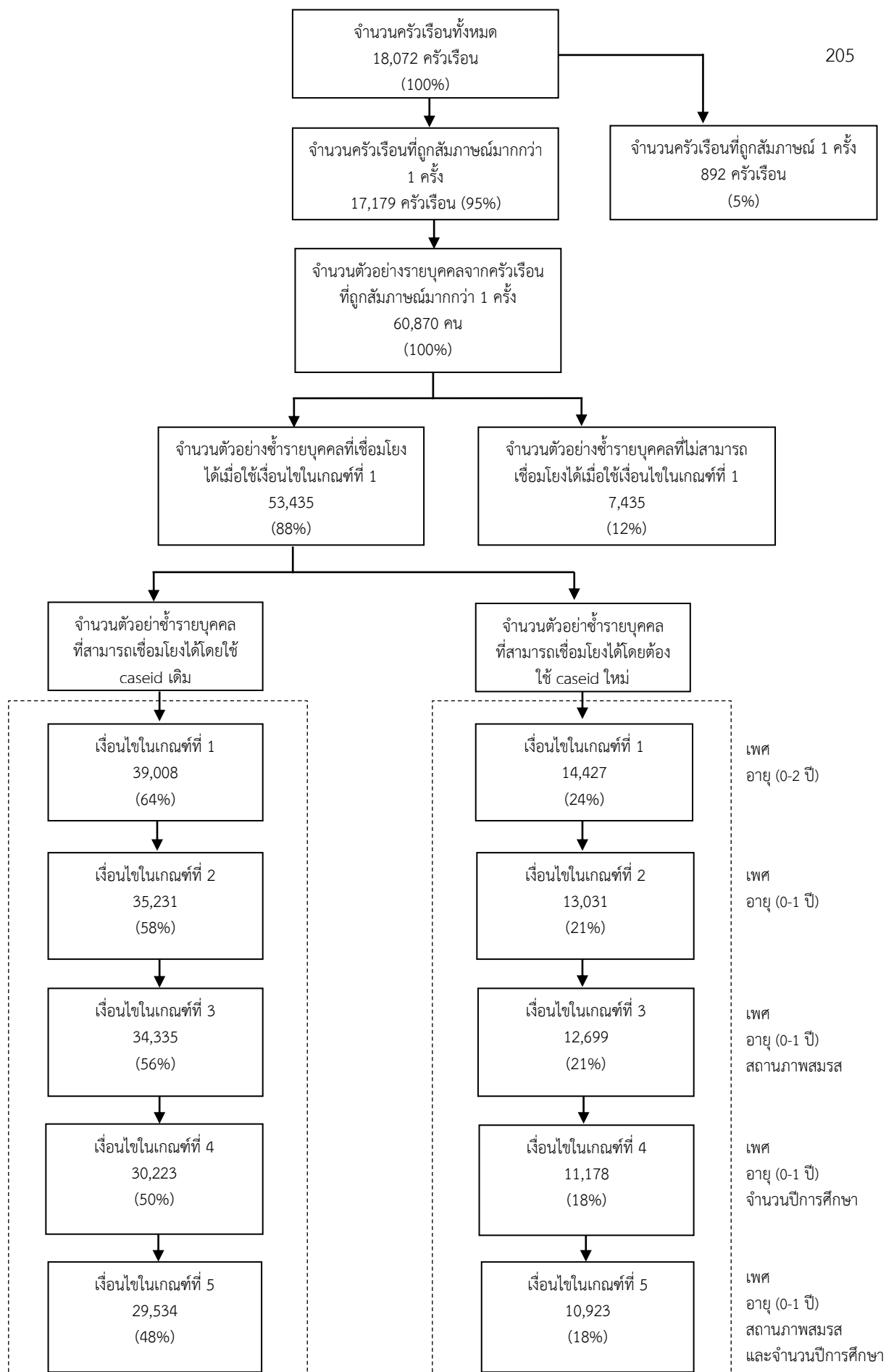
จากตารางที่ 3 พบว่า ครัวเรือนที่ 2 สมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 1 มีอายุ 42 ปี เพศชาย เมื่อสมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 2 มีอายุ 42 ปี เพศชาย และสมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 3 มีอายุ 35 ปี เพศชาย และสมาชิกคนที่ 1 ถูกสัมภาษณ์ครั้งที่ 4 มีอายุ 37 ปี เพศชาย ซึ่งจะเห็นว่า สมาชิกคนที่ 1 ที่ถูกสัมภาษณ์ครั้งที่ 3-4 ไม่ใช่คนเดียวกันกับสมาชิกคนที่ 1 ที่ถูกสัมภาษณ์ครั้งที่ 1-2 ดังนั้น เราจึงต้องสร้าง caseid ใหม่ให้กับสมาชิกคนดังกล่าว และเนื่องจากว่า สมาชิกที่ถูกสัมภาษณ์ครั้งที่ 3 มีอายุ 35 ปี และถูกสัมภาษณ์ครั้งที่ 4 มีอายุ 37 ปี เป็นไปตามเงื่อนไขในการสร้าง caseid_new เพราะมีอายุห่างกันไม่เกิน 2 ปี แต่ถ้าต้องสร้าง caseid_new2 ซึ่งมีเงื่อนไขว่าต้องมีอายุห่างกันไม่เกิน 1 ปี ทำให้สมาชิกคนนี้จะเป็นคนเดียวกันเมื่อใช้ caseid_new แต่จะเป็นคนละคนกันถ้าใช้ caseid_new2

ผู้ใช้สามารถเลือกใช้เกณฑ์การจัดทำกลุ่มตัวอย่างซ้ำได้จากตัวแปร `same_member` ซึ่งมีรายละเอียด ดังนี้



จากเกณฑ์ข้างต้น สามารถสรุปจำนวนข้อมูลกลุ่มตัวอย่างซ้ำรายบุคคลเฉลี่ยต่อ 1 rotation plan โดยพบว่ามีจำนวนครัวเรือนทั้งหมดโดยเฉลี่ย 18,072 ครัวเรือน แบ่งเป็นจำนวนครัวเรือนที่ถูกสัมภาษณ์เพียง 1 ครั้ง เฉลี่ย 892 ครัวเรือน คิดเป็นร้อยละ 5 ของครัวเรือนทั้งหมด และจำนวนครัวเรือนที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง เฉลี่ย 17,179 ครัวเรือน คิดเป็นร้อยละ 95 ของครัวเรือนทั้งหมด โดยที่ครัวเรือนที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง มีจำนวนตัวอย่างรายบุคคลทั้งหมดโดยเฉลี่ย 60,870 คน ซึ่งคิดเป็นตัวอย่างรายบุคคลที่ถูกสัมภาษณ์มากกว่า 1 ครั้งขึ้นไป มีจำนวนเฉลี่ย 53,887 คน คิดเป็นร้อยละ 89 ของจำนวนคนทั้งหมด (ใช้เงื่อนไขอายุห่างกันไม่เกิน 2 ปี) และตัวอย่างรายบุคคลที่ถูกสัมภาษณ์มากกว่า 1 ครั้งขึ้นไป มีจำนวนเฉลี่ย 53,154 คน คิดเป็นร้อยละ 87 ของจำนวนคนทั้งหมด (ใช้เงื่อนไขอายุห่างกันไม่เกิน 1 ปี)

เมื่อใช้เกณฑ์การค้นหาข้อมูลตัวอย่างซ้ำรายบุคคลที่อยู่ในครัวเรือนตัวอย่างซ้ำ (ครัวเรือนที่ถูกสัมภาษณ์มากกว่า 1 ครั้ง) พบว่า หากใช้เกณฑ์ที่ 5 ซึ่งรัดกุมที่สุดจะมีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ย 40,457 คน คิดเป็นร้อยละ 66 ของจำนวนคนทั้งหมด และหากใช้เกณฑ์ที่ 3 พบว่า มีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ย 41,401 คน คิดเป็นร้อยละ 68 ของจำนวนคนทั้งหมด และหากใช้เกณฑ์ที่ 4 พบว่า มีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ย 47,034 คน คิดเป็นร้อยละ 77 ของจำนวนคนทั้งหมด ในขณะที่หากใช้เกณฑ์ที่ 2 จะมีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ย 48,262 คน คิดเป็นร้อยละ 79 ของจำนวนคนทั้งหมด และเมื่อใช้เกณฑ์ที่ 1 ซึ่งมีความรัดกุมน้อยที่สุด จะพบว่า มีจำนวนตัวอย่างซ้ำรายบุคคลเฉลี่ย 53,435 คน คิดเป็นร้อยละ 88 ของจำนวนคนทั้งหมด ทั้งนี้ขึ้นอยู่กับผู้ใช้ข้อมูลที่ต้องตัดสินใจว่าจะเลือกใช้เกณฑ์ใด โดยจำนวนข้อมูลตัวอย่างซ้ำรายบุคคลสามารถแสดงได้จากรูปดังต่อไปนี้



แผนภาพสรุปจำนวนข้อมูลตัวอย่างข้าราชการเฉลี่ยต่อ 1 rotation plan