패널자료 품질개선 연구(VII)

www.kli.re.kr

김유빈 · 박민규 · 변종석 · 최효미 · 권혁진 · 김우영
성재민 · 이지은 · 신선옥 · 이혜정 · 김기홍

KLI
한국노동연구원

목 차

요 약 ····· i

제1장 서 론 ····· (김유빈) ····· 1

제2장 한국노동패널조사 표본 추가 연구 ·· (박민규 · 변종석) ····· 6

제1절 연구개요 ····· 6

1. 연구목적 ····· 6
2. 연구내용 및 방법 ····· 7

제2절 현행 노동패널 표본의 모집단 대표성 분석 ····· 9

1. 개 요 ····· 9
2. 현행 노동패널의 표본 가구원 분석 ····· 11
3. 현행 노동패널의 표본 가구 분석 ····· 20
4. 현행 노동패널의 가중치 분석 ····· 22

제3절 사례연구 : 해외 패널조사에서의 표본 보완 및 개편 사례 ····· 26

1. 미국의 PSID ····· 26
2. 영국 BHPS와 UKHLS ····· 29
3. 독일의 SOEP ····· 34
4. 호주의 HILDA ····· 38

제4절 한국노동패널조사(KLIPS)에서의 표본 확대 방안 ····· 41

1. 표본 확대 및 패널 관리 방안 ····· 41
2. 표본 확대에 따른 KLIPS 결합표본의 모집단과 가중치 ····· 56

제5절 추가 표본추출을 위한 2017년 시범조사 표본설계 ····· 85

1. 2015 인구주택총조사의 특징 ····· 86
2. 2017통합패널의 모집단과 표본틀 ····· 87
3. 노동패널 현황 ····· 108
4. 시범조사를 위한 새로운 표본 설계 ····· 118
5. 가중치와 추정식 ····· 136

제3장 KILPS 보육관련 설문조사의 문제와 데이터 품질 제고 방안 ····· (최효미) ····· 166

제1절 문제 제기 ····· 166

제2절 설문의 구조 변화와 누락 사유 ····· 168

제3절 회고조사 결과 ····· 175

제4절 보육・교육서비스 이용 행태 변화 ····· 179
1. 보육・교육서비스 이용 비중 변화 ····· 179
2. 보육・교육비용 변화 추이 ····· 185
3. 보육・교육비용 결정요인 ····· 190

제5절 사교육비 부담과 관련된 문항에 대한 검토 ····· 193

제6절 소 결 ····· 195

제4장 소득세 추정 모형(Personal income tax calculator)의 필요성과 활용 방안 ····· (권혁진) ····· 199

제1절 서 론 ····· 199

제2절 소득세 추정모형의 필요성과 국내・외 사례 ····· 200
1. 소득세 추정모형의 의의와 필요성 ····· 200
2. 국내・외 사례 및 시사점 ····· 211

제3절 소득세 추정모형의 특징 ····· 229
1. 2015년 귀속 소득세액 추정 업데이트 결과 ····· 229
2. 조사세액과 추정세액 비교 ····· 233

제4절 소 결 ····· 240

제5장 한국노동패널조사와 경제활동인구조사에 나타난 노동시장 활동 비교 ····· (김우영) ····· 244

제1절 서 론 ····· 244

제2절 연도별 노동시장 활동 현황 ····· 245

1. 전체 근로자 ····· 246
2. 남성 근로자 ····· 250
3. 여성 근로자 ····· 253

제3절 월별 노동시장 활동 현황 ····· 256

1. 전체 근로자 ····· 256
2. 남성 근로자 ····· 261
3. 여성 근로자 ····· 265

제4절 KLIPS와 통계청의 격차에 대한 원인 탐색 ····· 269

제5절 요약 및 제언 ····· 271

제6장 한국노동패널조사 웹서베이 도입 가능성 검토를 위한 연구 ····· (성재민) ····· 301

제1절 서 론 ····· 301

제2절 웹서베이의 장단점 ····· 302

제3절 웹서베이의 단점을 극복할 대안으로서 혼합조사(mixed mode) ····· 303

제4절 조사된 데이터에서 나타날 수 있는 이슈와 대안 ····· 305

제5절 대규모 패널조사 웹서베이 도입 사례 ····· 306

1. The Age 55 Survey of the National Child Development Study(NCDS) ····· 306
2. Health and Retirement Study(HRS) ····· 309

3. 영국 Understanding Society ····· 313
4. 기 타 ····· 325
제6절 웹서베이 도입을 위한 파일럿 서베이 제안 ····· 327

제7장 웹을 활용한 조사의 효과에 대한 실증연구 :
노동패널 CAWI 실험 데이터의 분석
····· (이지은 · 신선옥 · 이혜정 · 김기홍) ····· 330
제1절 서 론 ····· 330
제2절 노동패널 CAWI 실험 설계 개요 ····· 331
제3절 CAWI의 효과에 관한 실증분석 ····· 334
1. 조사성공률, 응답시간, 가구원수 대비 성공 가구원수 ····· 334
2. 항목응답률 ····· 341
3. 분포의 차이 ····· 346
4. 항목무응답률 ····· 355
5. 설문유형 실험 분석 ····· 362
6. 직산업 코딩 검정 ····· 366
제4절 소 결 ····· 372

참고문헌 ····· 375

표 목 차

<표 1- 1> 패널자료 품질개선 연구(I~VI) 세부 연구주제 ····· 2

<표 2- 1> 원패널 및 통합패널의 유효 표본 가구원수 ····· 13
<표 2- 2> BHPS의 표본 추가 현황과 규모 ····· 30
<표 2- 3> SOEP 표본유형별 추가표본 규모 ····· 35
<표 2- 4> SOEP 혁신 표본(SOEP-IS) 표본규모 ····· 38
<표 2- 5> HILDA 11차(2011년)의 표본특성별 표본규모
(가구와 개인) ····· 40
<표 2- 6> 한국노동패널조사의 조사차수별 표본 현황 ····· 44
<표 2- 7> 지역별 98표본 원표본가구의 유지율 ····· 48
<표 2- 8> 2010년 인구주택총조사의 모집단과 추출틀의 포함
범위 ····· 57
<표 2- 9> 세종시 센서스 결과(2013년, 통계청) ····· 58
<표 2-10> 2010년 인구주택총조사의 일반조사구 추출틀 ····· 59
<표 2-11> 2010년 인구주택총조사의 아파트조사구 추출틀 ····· 60
<표 2-12> 신축아파트 추출틀(2010년 인구주택총조사 이후
신축아파트) ····· 63
<표 2-13> 신축아파트를 결합한 2010년 인구주택총조사 기반의
추출틀 특성 ····· 64
<표 2-14> 2015년 인구주택총조사의 등록센서스 결과 ····· 65
<표 2-15> 2015년 인구주택총조사 가구 현황 ····· 66
<표 2-16> 2015년 인구주택총조사 인구 현황 ····· 67
<표 2-17> 2010년과 2015년 모집단 비교 ····· 68
<표 2-18> 가구주택기초조사의 조사구 규모(2014년 12월 기준) ····· 70
<표 2-19> HILDA의 벤치마크 기준 ····· 75

<표 2-20> 11차(2011년)조사에서 결합표본의 표본유형별 가구수 … 80
<표 2-21> 모집단 규모(광역시도별, 동/읍면부별 및 주택유형별 가구수) … 88
<표 2-22> 광역시도별, 동/읍면부별 및 가구원수별 가구수 … 90
<표 2-23> 2015년 인구주택조사구의 조사구 기준 가구수 및 인구수 특성 … 93
<표 2-24> 2015년 인구주택조사구의 가구수 규모별 분포 … 93
<표 2-25> 2015년 인구주택조사구의 인구수 규모별 분포 … 93
<표 2-26> 2015년 인구주택조사구의 가구수 분포별 가구수 및 인구수 특성 … 94
<표 2-27> 2015년 인구주택조사구의 인구수 분포별 가구수 및 인구수 특성 … 94
<표 2-28> 2015년 인구주택조사구의 광역시도별 가구수 및 인구수 특성 … 95
<표 2-29> 2015년 인구주택조사구의 병합조사구 기준 가구수 및 인구수 특성 … 98
<표 2-30> 2015년 인구주택조사구의 병합조사구 기준 가구수 규모별 분포 … 98
<표 2-31> 2015년 인구주택조사구의 병합조사구 기준 인구수 규모별 분포 … 98
<표 2-32> 2015년 인구주택조사구의 병합조사구 기준 가구수 규모별 가구수 및 인구수 특성 … 99
<표 2-33> 2015년 인구주택조사구의 병합조사구 기준 인구수 규모별 가구수 및 인구수 특성 … 99
<표 2-34> 2015년 인구주택조사구의 병합조사구 기준 광역시도별 가구수 및 인구수 특성 … 100
<표 2-35> 2015년 인구주택조사구의 병합조사구 기준 조사구 특성별 가구수 및 인구수 특성 … 102
<표 2-36> 2015년 인구주택조사구의 병합조사구 기준 동/읍면부별 가구수 및 인구수 특성 … 102

<표 2-37> 2015년 인구주택조사구의 병합조사구 기준 동/읍면부 및 조사구 특성별 가구수 및 인구수 특성 ····· 103
<표 2-38> 2015년 인구주택조사구의 병합조사구 기준 시도 및 동/읍면부별 가구수 및 인구수 ····· 103
<표 2-39> 2015년 인구주택조사구의 병합조사구 기준 시도 및 조사구 특성별 가구수 및 인구수 ····· 105
<표 2-40> 19차 노동패널의 광역시도별, 동/읍면부별 및 주택유형별 가구수 ····· 108
<표 2-41> 광역시도별, 동/읍면부별 및 가구원수별 가구수 ····· 111
<표 2-42> 광역시도별, 동/읍면부별 및 주택유형별 가구수 구성비 비교 ····· 113
<표 2-43> 광역시도별, 동/읍면부별 및 가구원수별 가구수 구성비 비교 ····· 115
<표 2-44> 광역시도별, 동/읍면부별 및 주택유형별 예상 상대표준 오차 ····· 120
<표 2-45> 17통합패널의 층화 ····· 121
<표 2-46> 17통합패널의 배분 방안별 표본 규모 ····· 122
<표 2-47> 통합패널 규모 및 과부족 표본 규모 ····· 126
<표 2-48> 통합패널 추가 조사구수 및 가구수 규모 ····· 128
<표 2-49> 추출틀의 모집단, 통합패널규모, 현행 패널 규모 및 추가 추출 규모(최종) ····· 130
<표 2-50> 통합패널의 최대허용오차 및 통합패널과 현행 패널의 예상 상대표준오차 비교 ····· 132

<표 3- 1> 미취학아동의 교육·보육 서비스 이용 설문 구조의 변화 ····· 169
<표 3- 2> 교육·보육 서비스 종류 문항의 용어 및 설문 수정 고려 사항(19차 설문 기준) ····· 170
<표 3- 3> 교육·보육비용 지원 정책 관련 설문 문항 및 개선점 (16차 이후 추가 문항) ····· 172

<표 3- 4> 코드값 변환 ····· 174
<표 3- 5> 회고조사 자료 ····· 176
<표 3- 6> 연도(조사차수)별, 연령별 분석 대상자수 ····· 177
<표 3- 7> 분석 대상자 특성(보육연령 5세만) ····· 177
<표 3- 8> 분석대상자 가구 특성(보육연령 5세만) ····· 178
<표 3- 9> 연도(조사차수)별 자녀 보육연령-출생연도 및 비용 지원 정책 대상 색인 ····· 179
<표 3-10> 연도(조사차수)별, 연령별 개별돌봄서비스 이용 아동 비중 ····· 182
<표 3-11> 출생연도-보육연령별 어린이집+유치원 이용 비중 ····· 183
<표 3-12> 연도(조사차수)별 보육・교육비용 미지출 가구 ····· 186
<표 3-13> 연도(조사차수)별 어린이집 월평균 이용비용 (어린이집 이용자) ····· 187
<표 3-14> 연도(조사차수)별 유치원 월평균 이용비용 (유치원 이용자) ····· 187
<표 3-15> 출생연도-보육연령별 어린이집 이용 비용 ····· 188
<표 3-16> 출생연도-보육연령별 교육·보육서비스 유치원 이용 비용 ····· 189
<표 3-17> 영유아 교육・보육비용 지출 결정요인 ····· 192
<표 3-18> 영유아 자녀 사교육비 관련 문항 응답이 차이를 보이는 경우(원자료 vs 회고자료) ····· 194

<표 4- 1> 정책목적별 소득공제의 분류 ····· 206
<표 4- 2> 근로소득세 납세자 수와 비중 ····· 230
<표 4- 3> 근로소득금액 구간별 평균 결정세액 : 최근 3년 ····· 231
<표 4- 4> 종합소득금액 구간별 평균 결정세액 : 최근 3년 ····· 232
<표 4- 5> 비교대상자 선정 ····· 234
<표 4- 6> 소득금액 구간별 납세 비중 비교:추정 vs. 조사 ····· 235
<표 4- 7> 소득금액 구간별 평균 세액 비교:추정 vs. 조사 ····· 237

<표 5- 1> KLIPS와 통계청 기준 경제활동지표(OECD 기준) ····· 247
<표 5- 2> 남성 KLIPS와 통계청 기준 경제활동지표 (OECD 기준) ····· 250
<표 5- 3> OECD 기준 KLIPS, 통계청 여성인구 총괄 ····· 253
<표 5- 4> 2009~2016년 월별 조사 항목 및 횟수 - 전체 근로자 ···· 257
<표 5- 5> 2009~2016년 KLIPS/통계청 비율(월평균) - 전체 근로자 ····· 260
<표 5- 6> 2009~2016년 월별 조사 항목 및 횟수 - 남성 ····· 261
<표 5- 7> 2009~2016년 KLIPS/통계청 비율(월평균) - 남성 근로자 ····· 264
<표 5- 8> 2009~2016년 월별 조사 항목 및 횟수 - 여성 ····· 266
<표 5- 9> 2009~2016년 KLIPS/통계청 비율(월평균) - 여성 근로자 ····· 269
<표 5-10> KLIPS와 통계청 기준 경제활동지표(OECD 기준) ····· 270
<표 5-11> 2개년 패널을 통한 동일 달, 다른 달 응답률 ····· 271

<표 6- 1> HRS의 파일럿 서베이 응답률 ····· 310
<표 6- 2> HRS의 파일럿 서베이 응답자들에 대한 설문 경험 피드백 조사 ····· 312
<표 6- 3> 5차 혁신패널 평가 결과 1 ····· 316
<표 6- 4> 5차 혁신패널 평가 결과 2 ····· 317
<표 6- 5> 5차 혁신패널 평가 결과 3 ····· 318
<표 6- 6> 5차 혁신패널 평가 결과 4 ····· 319
<표 6- 7> 6차 혁신패널 웹서베이 항목 무응답 실험 평가 결과 ···· 323

<표 7- 1> 조사도구별 접촉가구 대비 조사시작률 차이 (가구 설문) ····· 334
<표 7- 2> 조사도구별 조사성공률 차이 ····· 335
<표 7- 3> 조사도구별 접촉가구 대비 조사성공률 차이 (가구 설문) ····· 335

<표 7- 4> 대상가구원수 대비 응답가구원수 ····· 336
<표 7- 5> 가구별 대상가구원수 대비 응답가구원수 비율 차이 ····· 337
<표 7- 6> 조사도구별 대상가구원 전원 조사 기준 조사성공률 차이 ····· 337
<표 7- 7> 조사도구별 응답일수(가구 설문) ····· 338
<표 7- 8> 조사도구별 응답일수 차이(가구 설문) ····· 338
<표 7- 9> 조사도구별 응답시간 차이 ····· 338
<표 7-10> 조사도구별 연락 횟수(가구 설문) ····· 339
<표 7-11> 조사도구별 대리응답 분포 차이 ····· 340
<표 7-12> 조사도구별, 성별 교육연수 차이 ····· 340
<표 7-13> 조사도구별 연령 차이 ····· 341
<표 7-14> 가구 근로소득액 응답률(CAPI vs CAWI) ····· 341
<표 7-15> 가구소득 응답률의 집단(CAPI vs CAWI) 간 차이 ····· 342
<표 7-16> 금융자산 응답률의 집단(CAPI vs CAWI) 간 차이 ····· 343
<표 7-17> 지난 한 달 소득 응답률의 집단(CAPI vs CAWI) 간 차이 ····· 343
<표 7-18> 월평균 저축액 응답률의 집단(CAPI vs CAWI) 간 차이 ····· 344
<표 7-19> 가구 생활비 응답률의 집단(CAPI vs CAWI) 간 차이 ····· 344
<표 7-20> 주거지 정보 응답률의 집단(CAPI vs CAWI) 간 차이 ····· 345
<표 7-21> 임금근로자 비중 조사방법에 따른 차이 비교 ····· 345
<표 7-22> 종업원수 응답률의 조사방법에 따른 차이 비교 ····· 345
<표 7-23> 흡연 및 음주 응답률의 집단 간 차이 ····· 346
<표 7-24> 민감 문항 응답률의 집단 간 차이 ····· 347
<표 7-25> 입주·주거형태 및 가구원수 차이 검증 ····· 347
<표 7-26> 소득 변수 차이 검증 ····· 348
<표 7-27> 금융자산 및 기타변수 차이 검증 ····· 349
<표 7-28> 직장만족도 응답 차이 검증 ····· 350

<표 7-29> 생활만족도 차이 검증 ····· 351
<표 7-30> 건강상태 및 장애 응답 차이 검증 ····· 351
<표 7-31> 흡연 및 음주 변수의 차이 검증 ····· 352
<표 7-32> 종사상 지위 분포의 조사방법에 따른 차이 비교 ····· 353
<표 7-33> 임금근로자 중 비정규직 근로자 비중의 조사방법에 따른 차이 비교 ····· 354
<표 7-34> 평균근로시간의 조사방법에 따른 차이 비교 ····· 354
<표 7-35> 평균소득 유무의 조사방법에 따른 차이 비교 ····· 355
<표 7-36> 평균소득의 조사방법에 따른 차이 비교 ····· 355
<표 7-37> 무응답이 있는 문항의 무응답 비율 ····· 358
<표 7-38> 가구 생활비 무응답률의 집단(CAPI vs CAWI) 간 차이 ····· 358
<표 7-39> 부채(잔액과 상환금) 무응답률의 집단(CAPI vs CAWI) 간 차이 ····· 358
<표 7-40> 재학 여부, 혼인상태 무응답률(CAPI vs CAWI) ····· 359
<표 7-41> 작년 한 해 근로소득 무응답률의 집단 간 차이 ····· 360
<표 7-42> 일자리 시작시기(연/월/일) 무응답률의 집단 간 비교 ····· 361
<표 7-43> 임금근로자 중 비정규직 여부와 종업원수 무응답률의 조사방법에 따른 차이 비교 ····· 361
<표 7-44> 평균소득 무응답률의 조사방법에 따른 차이 비교 ····· 362
<표 7-45> 금융소득-응답률의 집단(A : B) 간 차이 분석 ····· 364
<표 7-46> 금융소득-분포 차이 검정 ····· 364
<표 7-47> (지난 한 달) 소득-응답률의 집단(A : C) 간 차이 분석 ····· 365
<표 7-48> (지난 한 달) 소득-분포 차이 검정 ····· 365
<표 7-49> 금융자산-응답률의 집단(B : C) 간 차이 분석 ····· 365
<표 7-50> 금융자산-분포 차이 검정 ····· 366
<표 7-51> 산업코드의 코더 간 불일치율 ····· 367
<표 7-52> 직업코드의 코더 간 불일치율 ····· 367
<표 7-53> 임금근로자의 산업코드 분포(코더 1) ····· 368

<표 7-54> 임금근로자의 산업코드 분포(코더 2) ····· 368
<표 7-55> 임금근로자의 직업코드 분포(코더 1) ····· 369
<표 7-56> 임금근로자의 직업코드 분포(코더 2) ····· 369
<표 7-57> 비임금근로자의 산업코드 분포(코더 1) ····· 370
<표 7-58> 비임금근로자의 산업코드 분포(코더 2) ····· 370
<표 7-59> 비임금근로자의 직업코드 분포(코더 1) ····· 371
<표 7-60> 비임금근로자의 직업코드 분포(코더 2) ····· 371

그림목차

[그림 2- 1] 패널조사의 구조 ····· 11
[그림 2- 2] 원패널 및 통합패널의 유효 표본 가구원수 ····· 12
[그림 2- 3] 성*연령에 대한 추계자료와 패널자료의 비교 ····· 14
[그림 2- 4] 가구 유형에 대한 추계자료와 패널자료의 비교 ····· 21
[그림 2- 5] 가중치 분포 ····· 23
[그림 2- 6] KLIPS의 결합표본의 결합 모집단 구조와 유형 ····· 54
[그림 2- 7] HILDA의 1차조사 횡단가중치 산출 과정 ····· 74
[그림 2- 8] HILDA의 종단가중치 산출 과정 ····· 77
[그림 2- 9] 결합표본의 가중치 산출 과정 ····· 79
[그림 2-10] 추가표본의 모집단 구조와 유형 ····· 80

[그림 3- 1] 연도(조사차수)별, 연령별 어린이집 이용 아동 비중 ····· 181
[그림 3- 2] 연도(조사차수)별, 연령별 유치원 이용 아동 비중 ····· 182
[그림 3- 3] 출생연도-연령별 어린이집 혹은 유치원 이용 아동 비중 ····· 184
[그림 3- 4] 국내 어린이집 및 유치원 이용률(2001~2015) ····· 185

[그림 4- 1] 우리나라 종합소득과세체계 ····· 203
[그림 4- 2] 근로소득자의 연말정산 세액계산 흐름도 ····· 204
[그림 4- 3] 소득증빙서류 수집률(2차~8차) ····· 222

[그림 5- 1] 경제활동참가율 ····· 248
[그림 5- 2] 고용률 ····· 248
[그림 5- 3] 실업률 ····· 249
[그림 5- 4] 남성 경제활동참가율 ····· 251

[그림 5- 5] 남성 고용률 ····· 252
[그림 5- 6] 남성 실업률 ····· 252
[그림 5- 7] 여성 경제활동참가율 ····· 254
[그림 5- 8] 여성 고용률 ····· 255
[그림 5- 9] 여성 실업률 ····· 255
[그림 5-10] 2009~2016년 월별 조사 횟수 - 전체 근로자 ····· 257
[그림 5-11] 2009~2016년 KLIPS와 통계청의 경제활동참가율 -
전체 근로자 ····· 258
[그림 5-12] 2009~2016년 KLIPS와 통계청의 고용률 - 전체 근로자 ·· 258
[그림 5-13] 2009~2016년 KLIPS와 통계청의 실업률 - 전체 근로자 ·· 259
[그림 5-14] 2009~2016년 월별 조사 횟수 - 남성 ····· 262
[그림 5-15] 2009~2016년 남성 KLIPS, 통계청 경제활동참가율 ····· 263
[그림 5-16] 2009~2016년 남성 KLIPS, 통계청 고용률 ····· 263
[그림 5-17] 2009~2016년 남성 KLIPS, 통계청 실업률 ····· 264
[그림 5-18] 2009~2016년 월별 조사 횟수 - 여성 ····· 265
[그림 5-19] 2009~2016년 여성 KLIPS, 통계청 경제활동참가율 ····· 266
[그림 5-20] 2009~2016년 여성 KLIPS, 통계청 고용률 ····· 267
[그림 5-21] 2009~2016년 여성 KLIPS, 통계청 실업률 ····· 268

[그림 6- 1] NCDS의 CAWI 응답률 추이 ····· 308
[그림 6- 2] 6차 혁신패널의 웹서베이 실험 조사 단계 구성 ····· 320
[그림 6- 3] 6차 혁신패널의 웹서베이 조사 화면 ····· 321
[그림 6- 4] 6차 혁신패널의 웹서베이 항목 무응답 실험 설계 ····· 322
[그림 6- 5] 8차 혁신패널의 웹서베이 실험 조사 단계 구성 ····· 323
[그림 6- 6] 8차 혁신패널의 웹서베이 유인 실험 설계 결과 ····· 324

[그림 7- 1] 표본 추출 과정 ····· 333
[그림 7- 2] 설문유형 A, B, C에 따라 화면에 보여주는 방식 ····· 363

요 약

1. 한국노동패널조사 표본 추가 연구

한국노동패널조사는 2009년의 표본 추가에도 불구하고 패널 마모로 인하여 전체 표본 가구의 수가 매해 지속적으로 감소하고 있다. 또한 지속적으로 진행되고 있는 모집단 변화를 현재 패널 표본이 온전히 대표하지 못하고 있는 것으로 판단되고, 패널자료의 적극적인 정책적 활용을 위하여 요구되는 정책 대상별 통계 생산 및 대응되는 부모집단의 종단면 분석을 위해서는 충분한 수의 표본이 확보되어야 한다. 이에 본 연구는 한국노동패널에 대한 정책적 활용에 대한 요구와 모집단의 변화를 반영하기 위하여, 현행 패널에 대한 분석 및 이에 근거한 패널 추가에 대한 필요성에 대한 연구, 적정한 수준의 추가표본 규모에 대한 연구, 그리고 실제 표본설계에 대한 사전연구를 수행한다.

한국노동패널조사의 2017통합패널 표본추출틀은 2015년 인구주택총조사의 일반가구 조사구이며, 일반가구 조사구는 특수조사구와 섬조사구를 제외한 단독주택, 연립주택 및 다세대주택 등의 보통조사구와 아파트조사구로 구성된다. 2017통합패널 구축을 위한 표본추출틀의 규모를 보면, 일반가구 중 특수조사구와 섬조사구 가구를 제외하고 표본추출틀에 포함된 가구수는 19,100,262가구이다.

2017통합패널을 위한 새로운 표본설계에서는 2015년 인구주택총조사를 모집단 및 조사구로 선정하여 광역시도별 동읍면부 및 주택유형별 표본을 배분한 후, 1998년 및 2009년 패널에 대한 연속성을 유지하도록 패널을 구축해야 하므로 2017통합패널에서는 층별로 현행 패널을 우선 확정하고 층별로 부족한 패널을 추가로 추출한다. 즉

17통합패널은 기본적으로 1998년 혹은 2009년 구축된 패널 구축을 위한 표본 설계의 틀을 유지하고, 2015년 인구주택총조사에서 나타난 광역시도별 동읍면부 및 주택유형별 변화된 모집단 특성과 현행 패널에서 추적 실패 및 패널 탈락 등으로 인해 마모된 패널을 보완하도록 표본설계를 수행한다.

17통합패널의 가중치 산출 방안은 다음 두 가지의 산출 방안이 가능하다. 첫째, 17년 조사에 참여한 모든 응답 가구 및 가구원만을 17통합패널의 원패널로 정의해, 기존의 각 패널의 추출 과정을 무시하고 원패널 중 응답 완료된 패널에 대해서만 사후층화 가중치로 17년 횡단가중치를 산출하는 방안(방안 1), 둘째, 17통합패널은 1998년 원패널, 2009년 추가 패널 및 2017년 추가 패널로 구성되므로 패널 추출과정으로 각각 고려한 패널가중치를 별도로 산출하여 통합패널을 위한 통합가중치를 산출하는 방안(방안 2). 통합패널의 가중치 산정은 시범조사의 진행 과정 전후 사항을 고려하여 추후 결정하도록 한다.

2. KILPS 보육관련 설문조사의 문제와 데이터 품질 제고 방안

한국노동패널조사는 교육・보육서비스 이용 실태에 관한 설문구조와 문항을 수정하는 과정에서 보육연령에 대한 CAPI 로직상의 오류가 발생했던 것으로 판단된다. 16차년도(2013년) 조사 이후 보육연령 5세(취학 전 1년)의 어린이집 및 유치원 이용 실태 및 이용비용 문항에 대한 조사가 누락되는 현상이다. 즉 결론적으로 16차년도(2013년 조사)~18차년도(2015년조사) 자료는 보육연령 5세(취학 전 1년)의 어린이집 및 유치원 이용 실태 및 이용비용이 전체적으로 누락되었으며, 19차년도 조사 초기 응답 가구 중 일부에서 응답 값이 없는 상황이 발생하였다. 따라서 본 연구는 노동패널 16차~19차 자료의 영유아 보육・교육서비스 회고조사 자료의 특징과 이를 원자료에 포함시켜 분석할 경우 발생할 수 있는 오류를 사전에 진단하고, 응답 값의

신뢰도를 검토하는 것을 주요 목적으로 한다. 다른 연령 및 다른 차수와의 비교, 종단자료로서 코호트 간의 비교 등을 통해 전반적인 추이에 이상한 변화가 섞여 있지는 않은지를 중점적으로 검증한다.

본 연구의 시사점은 다음과 같다. 회고조사 자료와 기존 자료를 통해 살펴본 보육연령 5세의 보육·교육서비스 이용 및 비용 변화 추세를 살펴보면, 실제 16차년도(2013년)에는 상당한 정도의 변화가 포착되었다. 노동패널 16차년도(2013년) 회고조사 자료는 이전 연도와의 추세를 비교해 볼 때 어린이집 혹은 유치원 이용이 급격히 증가하는 등의 변화가 관측되지만, 이러한 현상은 당해연도 조사가 완료된 만 3~4세에게서도 동일하게 관측되는 현상일 뿐 아니라 정책 변화로 인해 이용 행태가 크게 변동한 전년도(2012년) 자료와 유사한 패턴을 보이는 현상이므로, 16차년도~18차년도에 걸친 노동패널 보육연령 5세에 대한 회고조사 자료는 전반적으로 조사 과정과 결과에 있어 큰 문제가 없다고 판단된다. 하지만, 16차년도 이후 노동패널 보육·교육서비스 관련 자료는 회고조사로 인한 문제보다는 설문구조의 변경에 따른 문제가 오히려 큰 것으로 보이며, 향후에는 이에 대한 개선이 반드시 요구된다. 보다 세분된 서비스 종류에 대한 분석을 희망하는 연구자들을 위해서는 기존 데이터 구조를 가진 자료를 패치파일로 제공하고, 이후로는 16차년도 설문 변경의 내용을 보다 잘 반영할 수 있는 데이터 구조를 결정하여 원자료와 함께 제공하는 방안에 대한 전향적 검토가 요구된다. 기존 방식으로 원자료를 재구성하여 통합 데이터를 제공하는 경우라 할지라도 회고조사로 진행된 자료에 대해서는 회고조사 자료임을 명확하게 알 수 있도록 '회고조사 여부' 변수를 신설하여 함께 제공하는 것을 검토할 필요가 있다.

3. 소득세 추정 모형(Personal income tax calculator)의 의의와 활용 방안

소득세추정모형(Personal income tax calculator)은 미시 횡단면

혹은 패널 자료를 주로 활용하며, 기본적으로 현행 소득세제하에서 개인·가구들이 부담하는 납세액을 산출(calculate)한다. 이러한 맥락에서 많은 국가와 연구소에서 소득세추정모형(tax calculator) 혹은 조세/복지급여 모형(tax/benefit model)들을 개발·활용하고 있으나, 우리나라에서는 지속적으로 유지·활용되고 있는 모형이 아직까지 존재하지 않는 상황이다. 이에 본 연구는 노동패널 자료를 활용하여 소득세추정모형(tax calculator)을 개발·구축하고, 나아가 이를 많은 연구자가 활용하기 위한 방안을 검토하고자 한다.

2007~2014년 기간 동안의 소득세액 추정을 2015년 귀속 소득세액까지 업데이트한 결과를 살펴보면, 과세대상자에 대한 추정결과가 국세통계에 비해 약 100만~200만 명 정도 높게 나타난다. 노동패널이 다른 패널조사 자료들과 마찬가지로 고소득 계층에 대한 표본이 상대적으로 과소하다는 점을 고려하면, 저소득 계층에 대한 과대추정 정도가 상대적으로 더 클 것으로 예상되나, 근로소득금액이 1천만 원 이하 계층의 경우는 국세통계에 비해 추정 규모가 상대적으로 적게 추정된다는 점에 유의할 필요가 있다.

납세자 비중 역시 국세통계에 비해 추정결과가 높게 나타나고 있다. 다양한 소득공제 항목들을 비롯해 세액 계산에 필요한 정보들이 부족하기 때문에 분석대상 개인들의 세액이 높게 추정될 가능성이 높고, 이는 납세자 비중을 높이는 요인이다.

근로소득금액 구간별 납세자수와 평균 세액에 대한 추정치를 국세통계와 비교하면, 소득금액 구간별 납세자수에 대한 추정결과를 전반적으로 보면, 저소득계층의 납세자수는 국세통계에 비해 많게 추정되고 있는 반면, 소득수준이 높아질수록 상대적으로 적은 규모로 추정된다. 그러나 소득금액이 1천만 원 이하 계층의 납세자 규모와 비중은 2014년 이전과 이후가 다소 상반된 결과가 나타난다. 즉 2014년 이전에는 국세통계에 비해 납세자 비중과 규모가 크게 추정된 반면, 이후에는 상대적으로 적게 추정되고 있다.

2015년의 경우 평균 근로세액에 대한 추정결과를 보면, 2014년 추

정결과와 유사하게 나타난다. 이는 2015년 당시 세제개혁 폭이 크지 않았던 점이 반영된 것으로 판단된다.

종합소득세액에 대한 추정결과와 국세통계를 비교하면, 종합소득세 추정치에 대한 비교 결과 역시 2014년과 2015년이 유사하고, 2013년과 비교할 때 2014년 이후 세부담은 저소득계층은 축소, 고소득계층은 증가라는 패턴을 확인할 수 있다. 하지만, 1천만 원 이하 계층에 대한 평균 세액 추정 결과를 보면, 2014년 이후부터는 국세통계에 비해 크게 추정된다는 점에 유의할 필요가 있다.

근로소득자에 대한 비교 결과를 보면 4천만 원 이상 구간은 대부분 납세자 비중이 100%에 가깝기 때문에 비교할 필요가 없으나, 4천만 원 이하 계층의 경우에는 조사 결과와 추정 결과 간 차이가 상당히 크다. 특히, 1천만 원 이하 구간에서 조사자료상 납세자 비중은 추정결과에 비해 상당히 높게 나타난다. 또한 국세통계상의 납세자 비중을 기준으로 보면 추정치가 국세통계 실적치와 유사한 것을 볼 때, 개인들의 응답이 지나치게 높다고 평가할 수 있다. 종합소득자의 경우는 근로소득자와는 정반대의 결과가 나타나, 추정결과에 비해 조사자료상의 결과가 상당히 낮게 나타나고 있다. 평균세액에 대한 비교는 근로소득세와 종합소득세 모두 유사하게 나타난다. 즉 4천만 원 이하 계층에서 조사세액이 지나치게 크게 나타나고 있고, 소득수준이 낮을수록 추정세액과 조사세액의 차이는 더욱 증가한다.

4. 한국노동패널과 경제활동인구조사에 나타난 노동시장 활동 비교

본 연구는 KLIPS의 통합표본이 시작되는 2009년 이후 한국노동연구원에서 공급하고 있는 한국노동패널(KLIPS)에서 측정되는 노동시장 활동 지표(경제활동참가율, 고용률, 실업률)와 통계청의 경제활동인구조사에서 나타나는 지표를 시계열적으로 비교하고, 이들의 차이정도를 확인하여 그 원인과 개선을 위한 제언을 제시하는 데 목적

을 둔다.

한국노동연구원의 KLIPS와 통계청의 경제활동인구조사에서 측정되는 경제활동참가율, 고용률, 실업률을 비교한 결과, KLIPS에서 얻어진 경제활동참가율과 고용률이 통계청 지표보다 전반적으로 낮게 측정되지만 그 오차는 그리 크지 않게 나타났다. 또한, 남성과 여성을 구분하여 비교하여도 패턴에 큰 차이를 보이지 않았다. 한편, 두 자료의 시계열 간 상관관계를 구한 결과, 경제활동참가율과 고용률은 양의 상관관계를 보이며, 특히 여성의 경우 상관관계가 더 강하게 나타났다. 하지만 실업률의 경우에는 상관관계가 없는 것으로 나타나기 때문에 KLIPS의 실업률이 통계청의 실업률을 대리하기에는 적절하지 않다고 판단된다. KLIPS에서 경제활동참가율, 고용률, 실업률을 월별로 측정한 결과, 조사빈도가 상대적으로 높은 5월에서 10월까지의 자료가 통계청 경활조사의 자료와 매우 근접한 것으로 나타나고 있다. 다만 이 경우에도 실업률은 상대적으로 오차가 크게 측정되고 있다. KLIPS의 경제활동참가율, 고용률이 낮은 것은 비경제활동인구가 너무 많이 표집되거나 실업자와의 구분이 명확지 않기 때문이다. 한편, KLIPS의 월별 통계는 연속성이 부족하기 때문에 이를 이용하여 월별 노동시장활동 통계를 계산하는 것은 지양해야 할 것이며, KLIPS의 월별 통계의 정확성을 높이기 위해서는 응답자의 응답시기를 가능한 한 동일한 달로 집중시킬 것을 제안한다.

5. 한국노동패널조사에 웹서베이 도입 가능성 검토를 위한 연구

본 연구의 주목적은 다음과 같다. 첫째, 본 연구는 한국노동패널조사에 웹서베이 도입 가능성을 검토하고자 한다. 우리나라 대규모 패널조사에서 웹서베이는 도입 경험이 없으므로, 본 연구에서는 왜 웹서베이 도입 논의가 필요한지부터 검토하도록 한다. 둘째, 본 연구는 웹서베이의 도입 가능성 검토를 위해 먼저 웹서베이가 갖는 장단점

을 문헌연구를 통해 요약한다. 셋째, 대규모 패널조사에서 웹서베이를 도입하거나 도입을 위해 실험이 이루어지고 있는 사례들을 검토한다.

웹서베이의 장점은 다음과 같다. 첫째, 비용이 적게 든다. 둘째, 좀 더 빠르게 조사를 진행할 수 있다. 셋째, 문항의 스킵(skip) 구조가 설문조사 응답을 어렵게 만드는 경우가 많은데, 웹서베이는 이미 프로그램 된 문항이 자동으로 뜨기 때문에 오류 가능성이 낮다. 필요하다면 프로그램상으로 임의 문항 할당 등이 자동으로 가능하기 때문에 설문조사에서 유연성을 확보할 수 있다는 장점도 있다. 측정(measurment)의 측면에서도 웹서베이는 몇 가지 장점이 있다. 우선, 응답자가 민감하게 생각할 수 있는 질문에서 자기기입식이라는 특징 때문에 프라이버시 보호가 가능하고, 이로 인해 좀 더 정직한 응답을 유도할 수 있다는 장점이 있다.

또 다른 큰 장점은 설문구성에서 음악, 비디오, 그림을 넣을 수 있어 질문의 한계를 넓혀준다는 점이다. 다만, 잘못 설계될 경우 5점 척도가 좋으냐 7점 척도가 좋으냐, 스케일을 표현하는 단어를 무엇으로 선택하느냐에 따라 응답결과가 달라지듯 오류를 야기할 가능성이 있다는 점은 주의해야 할 사항이다. 설문에 대한 적극성을 가진 응답자라면 설문의 속도를 스스로 통제하면서 좀 더 철저한 답변을 기대할 수 있다는 것도 장점이다.

그러나 웹서베이를 하면서 채팅을 하고 웹서핑을 하는 등 서베이에 대한 집중력을 해치는 행동이 함께 발생할 가능성도 높다는 단점도 동시에 존재한다. 그러나 웹서베이는 커버리지(coverage) 측면에서 중대한 약점이 있으며, 또한 다른 조사에 비해 응답률도 낮다고 한다. 이 외에도 면접원이 없으므로 일단 설문을 시작한다 하더라도 설문을 마쳐야 한다는 압박에서 자유로워 완성된 설문결과를 확보할 가능성이 낮을 수 있다는 점도 단점으로 지적될 수 있다.

혼합조사의 방식에 있어서는 다음과 같은 조합을 생각해 볼 수 있다.

첫째는, 다른 조사 방식으로 컨택하고, 응답방식은 하나의 방식으로 하는 혼합조사 방식이다. 둘째, 설문의 일부분만 다른 조사방식을 활용하는 방법이다. 셋째, 응답자 특성에 따라 다른 조사 방식을 제공하는 방법이다. 넷째, 응답자 설득이 매우 중요한 첫 번째 조사에서는 대면조사를 하고 두 번째 조사부터는 웹서베이를 하는 것도 대안이 될 수 있다.

웹조사 방식의 조사데이터에서 나타날 수 있는 이슈는 다음과 같다. 데이터조사 방식이 다를 때 조사된 결과가 다를 수 있다. 이에는 세 가지 이유가 있다(de Leeuw & Hox, 2011). 첫 번째, 각 조사방식에 잘 반응하는 표본의 특성이 다를 경우. 두 번째, 조사 방법마다 질문 포맷이 달라서 다른 응답결과가 야기되는 경우. 세 번째, 조사 방법 자체에 내재된 편향이 다른 응답으로 이끄는 경우이다. 이 중 첫 번째는 혼합조사 방법의 단점이 아니라 장점이다. 혼합조사의 원래적 목적인 커버리지 문제와 무응답 편향이 교정된 결과라 받아들여야 하기 때문이다. 두 번째와 세 번째가 문제이다. 두 번째 문제는 조사 방식마다 잘 조사되는 설문지 구성이 다를 수 있고, 개별 조사 방식에 최적화된 방식으로 설문지가 만들어졌을 경우 발생한다. 세 번째 문제는 순수한 의미의 mode effect이다. 이를테면 면접조사에서 흔히 나타나는 사회적 바람직함 편향이 그런 예가 된다. 가장 분명한 차이는 사회적 바람직함 편향이므로, CAWI와 CAPI를 이용한 혼합조사라면 사회적으로 민감한 문항은 별도의 설문지를 제공하든지, 면접원이나 가족 같은 다른 응답자가 보지 못하는 곳에서 응답자가 직접 입력하도록 하는 것이 대안이 될 수 있다.

6. 웹을 활용한 조사의 효과에 대한 실증연구

패널조사상의 비성공 가구들은 해마다 사망, 이민 등의 필연적 사유로 발생할 수밖에 없다. 비성공 사유를 살펴보면, 노동패널 초기에는 강력거절이 주를 이뤘으나, 최근 5년 동안은 이사추적실패가 주

요 원인으로 꼽히고 있다. 이는 패널이 조사를 거부하는 것이 아니라, 연락처가 없어서 찾아가지 못하는 것이기 때문에, 이를 해결한다면 패널 유지율을 더욱 안정적으로 가져갈 수 있을 것이라 기대된다.

웹서베이가 이에 대한 대안 중 하나가 될 수 있다. 본 연구에서는 웹서베이 실험연구를 실시하여, 자료의 체계적인 분석을 통해 웹서베이의 도입을 검토하고자 한다.

분석 결과는 다음과 같다. 첫째, CAWI는 CAPI에 비해 조사시작은 용이하지만, 마지막까지 조사를 완료하는 비율은 현저히 낮았다. 또한 가구에 속한 15세 이상 가구원 전수를 조사해야 함에도 불구하고, 응답을 거부하는 가구원이 많아 개인의 응답률이 저조하였고, 대리응답도 더 많은 것으로 나타났다. 성별이나 학력에 따라서는 차이를 보이지 않지만, CAWI 응답자의 연령이 더 낮은 것으로 조사되었다.

둘째, 항목응답률에서는 가구소득, 금융자산, 지난 한 달 근로소득, 월평균 저축액, 금주 등의 변수는 CAPI의 응답률이 높았다. 그러나 흡연, 금연 항목, 직업훈련, 사회보험수급여부 등의 변수는 CAWI의 응답률이 높게 나타났으며, 일자리 관련 변수는 차이가 없는 것으로 나타나, 변수 특성에 따라 응답률의 차이가 발생하였다.

셋째, 분포의 차이를 살펴보면, 가구소득변수는 근로소득액을 제외한 금융소득액, 부동산소득액, 이전소득액, 기타소득액, 개인의 평균 소득은 분포의 차이는 없었다. 다만 근로소득의 경우 t-test 검정에서 비모수 검정에서는 유의한 차이가 없는 것으로 나타났으며, 평균금액은 CAPI가 CAWI보다 더 높았다. 금융자산과 지난 한 달 소득, 월평균 저축액은 차이가 있는 것으로 조사되었다. 만족도 항목 중에서는 전반적인 일자리(직장)만족도, 전반적 일(직무)만족도, 전반적 생활만족도에서는 CAWI가 높았고, 세부 항목에서는 통계적으로 유의한 차이를 발견할 수 있었다. 흡연과 금연시도, 금주시도는 분포에서도 차이를 보였다. 일자리 관련 변수 중 고용형태와 평균근로시간은 통계적으로 유의한 차이를 보이지 않았지만, 통계적 기준

의 비정규직 여부에서 차이가 나타났다.

넷째, 항목 무응답률은 가구변수, 교육, 혼인상태, 일자리시작시기 등의 변수는 조사방법에 따라 차이가 없는 것으로 조사되었다. 그러나 세전 근로소득, 세후 근로소득, 종업원수는 CAWI의 무응답이 더 높은 것으로 나타났다.

다섯째, CAWI 방법 내에서 설문유형에 따른 차이는 금융소득, 지난 한 달 소득, 금융자산 등 세 가지 항목에 대해 모두 통계적으로 유의한 차이가 없어서 특정한 설문유형을 채택할 필요는 없는 것으로 나타났다.

여섯째, 직산업 코딩은 2명의 코더가 각자 작성하여 그 차이를 검정하였다. 대분류 기준으로 산업은 CAPI보다 CAWI에서 임금근로자, 비임금 근로자에서 모두 불일치가 높게 발생하였고, 직업은 비임금근로자에 한해 차이가 발생하였다.

이상의 분석결과를 볼 때, 조사성공률 측면에서 CAWI가 CAPI보다 효과적이지 않은 것으로 생각된다. 항목응답률, 분포의 차이, 항목 무응답률, 직산업 코딩 등의 검정을 통해서 일부 항목들에서는 조사 시행 전에 설문 문항의 보완이나 후속 전화조사의 보완이 필요한 것으로 보이나, 비성공 가구들에 대한 보완적인 조사방법 중의 하나로서 CAWI 접근은 검토될 필요성이 있다.

제 1 장
서 론

본 연구는 「패널자료 품질개선 연구 시리즈」의 7차 연구이다.

한국노동연구원 패널데이터연구실은 생산·품질제고·연구분석 등 전 과정에 걸쳐 전담 관리하는 한국노동패널조사(Korean Labor and Income Panel Studies : KLIPS)와 사업체패널조사(Workplace Panel Study : WPS)의 품질개선을 위해 2012년도 1차 연구를 시작으로 현재까지 7차에 걸쳐 패널자료 품질개선 연구를 수행하고 있다. 패널자료 품질개선 연구의 주제는 표본의 설계·추출, 가중치 조정, 표본이탈·무응답과 히핑 및 측정오차의 보정, 자료검증과 기존 항목을 이용한 변수추정 등 패널조사자료의 품질개선을 위한 다양한 세부 주제들로 구성되어 있다. 패널자료 품질개선 연구의 지난 차수별 세부 연구주제는 다음과 같다.

지난 패널자료 품질개선 연구 시리즈의 연장선에서 본 연구인 「패널자료 품질개선 연구 Ⅶ」에서는 한국노동패널조사 표본 추가 연구, 한국노동패널조사 보육관련 설문의 문제와 데이터 품질 제고 방안, 소득세 추정 모형의 의의와 활용 방안, 한국노동패널조사와 경제활동인구조사에 나타난 노동시장 활동 비교, 한국노동패널조사 웹서베이 도입 가능성 검토를 위한 연구, 웹을 활용한 조사의 효과에 대한 실증연구를 수행하였다.

2018년 20차 조사를 앞두고 있는 한국노동패널조사는 패널자료의 속성상 패널 마모로 인하여 표본수가 지속적으로 감소하고 있다. 한국노동패널은 정책 대상별 통계 생산 및 횡·종단면 분석 등 패널자료의 적극적인 정책적 활용을 위해 패널자료의 품질개선뿐만 아니라, 표본의 확충에도

〈표 1-1〉 패널자료 품질개선 연구(Ⅰ~Ⅵ) 세부 연구주제

패널자료 품질개선연구 (Ⅰ)
· 사업체패널자료 가중치조정 연구
패널자료 품질개선연구 (Ⅱ)
· 개인조사자료와 사업체조사자료 간 임금 분포의 히핑현상 비교
· 조사 방법과 히핑의 관계
· 노동패널 표본이탈 연구
· 사업체패널조사에 나타난 표본이탈 현황 및 대체 표본에 대한 분석
· 고령화연구패널조사 건강영역의 품질개선 방안
패널자료 품질개선연구 (Ⅲ)
· 소득 히핑의 보정
· 측정오차에 대한 통계적 보정 방법론 연구
· 표본이탈 교정을 위한 계량경제학 방법
· Nonignorable panel attrition에 대한 보정 연구
· Paradata를 활용한 패널 무응답 보정을 위한 가중치 부여 방안
패널자료 품질개선연구 (Ⅳ)
· 동태적 패널모형에서 표본이탈 교정
· Paradata를 이용한 무응답 자료 회귀분석
· 사업체 패널조사의 표본 설계 관련 연구
· 한국에서의 자료검증 연구의 필요성에 대하여
패널자료 품질개선연구 (Ⅴ)
· 사업체패널조사의 시점 및 표본 대표성
· 사업체패널조사의 표본이탈과 대표성
· 표본 이탈률의 크기와 편향의 관계
패널자료 품질개선연구 (Ⅵ)
· 사업체패널조사의 시점 및 표본 대표성
· 사업체패널조사의 표본이탈과 대표성
· Propensity 점수를 이용한 사업체 조사의 대표성 연구
· 표본 이탈률의 크기와 편향의 관계
· 2016 사업체패널조사의 추가표본설계 및 표본추출
· 노동패널 항목 무응답 처리방법 보완
· 사업체패널조사의 신뢰도와 타당도 검토
· 노동패널자료 소득세액 추정

노력을 기울여 왔다. 이와 같은 정책적 요구와 모집단 변화를 반영하기 위해 한국노동패널조사는 2009년의 1차 표본추가에 이어, 20차 조사부터 2차 표본추가를 계획 중에 있다. 패널의 표본 확충을 위해서는 사전적으로 현행 패널에 대한 분석 및 이에 근거한 패널 추가에 대한 필요성에 대한 연구, 적정한 수준의 추가 표본 규모에 대한 연구, 그리고 실제 표본설계에 대한 연구가 선행되어야 한다. 이에 본 연구의 제2장에서는 패널의 확대 및 대표성 확보에 대한 필요성을 검토하는 한편, 가능한 표본 추가 방안을 제안하고 이를 위해 필요한 표본설계의 기본적인 틀을 제안한다.

패널조사의 품질 제고를 위해서는 위에서 언급한 대표성의 유지 및 확보뿐만 아니라, 설문조사의 방식 및 항목별 응답에 대한 지속적 관리·검토가 요구된다. 한국노동패널조사는 가구 내 자녀에 대한 교육·보육서비스 이용 실태에 관한 문항을 포함하고 있다. 교육·보육서비스 이용에 관한 설문 문항은 매해 조사되는 본조사항목(2·6차년도 제외)이지만, 조사 문항과 설문구조의 변화가 수반되었다. 한국노동연구원 한국노동패널팀은 위의 과정에서 16차년도(2013년) 조사 이후 보육연령 5세의 어린이집 및 유치원 이용 실태 및 이용비용 문항에 대한 조사가 누락되는 현상을 발견하고, 이에 대한 대응으로 19차(2016년) 조사의 설문로직 개편과 더불어 16차년도(2013년)~18차년도(2015년) 조사 미응답자에 대해 보육연령 5세 당시 상황에 대한 회고조사를 실시했다. 이에 본 연구의 제3장에서는 한국노동패널 16~19차 조사의 영유아 보육·교육서비스 회고조사 자료의 특징과 이를 원자료에 포함시킬 시 발생할 수 있는 오류를 사전에 진단하고, 응답 값의 신뢰도 검토를 목적으로 연구를 수행하였다.

한국노동패널조사는 한국을 대표하는 종단면조사로서 한국 가구 및 가구원의 경제활동상태 및 노동시장 활동 전반에 관한 설문을 포함하고 있지만, 세부항목에 있어 모든 내용을 파악·포괄하는 것은 불가능하며, 소득세 관련 정보가 이에 해당한다. 소득세추정모형은 미시 횡단면 혹은 패널자료를 주로 활용하여 현행 소득세제하에서 개인·가구들이 부담하는 납세액을 산출하는 모형으로, 소득세추정모형의 개발과 활용은 과세가 개인·가구의 경제·사회적 행위에 미치는 영향에 대한 다양한 연구를 수행할 수 있게끔 한다는 점에서 유용하다. 이에 본 연구의 제4장에서는

한국노동패널조사 자료를 활용하여 개발하고 있는 소득세추정모형의 주요 결과를 제시하여 이의 활용성과 타당성을 검토한다.

패널조사를 포함하여 모든 조사의 활용성과 타당성은 조사 자료가 반영하고자 하는 현실의 대표성을 어느 정도 확보하느냐에 달려 있다. 이를 위해서는 가용한 각 항목별 조사자료 간의 비교를 통해 조사자료의 대표성을 지속적으로 검토하는 것이 중요하다. 이에 한국노동패널팀은 한국노동패널조사의 품질개선을 위한 조사자료 간 비교연구를 수행해 오고 있다. 본 연구의 제5장은 위의 연장선에서 한국노동패널조사와 경제활동인구조사의 고용지표(경제활동참가율, 고용률, 실업률)의 비교연구를 수행한다. 한국노동패널조사는 매년 가구원을 추적조사하는 종단면 조사이나, 매년 조사 시기를 특정하는 것은 불가능하다. 이런 측면에서 패널조사의 각 고용지표별, 월별 분석을 시행하는 것이 적절한지에 대한 의문이 남는 것이 사실이다. 이에 본 연구는 한국노동패널조사에서 측정되는 고용지표와 경제활동인구조사 간 차이를 시계열적으로 비교하는 한편, 격차의 원인을 파악하고, 지표의 개선을 위한 제언을 제시한다.

본 연구의 제6장과 제7장은 한국노동패널조사의 웹서베이 도입 가능성과 이의 활용성 및 효과에 대한 실증연구를 수행한다. 패널조사의 성격상 비성공 가구의 증가로 인한 패널 마모는 필연적으로 발생할 수밖에 없다. 이러한 비성공 사유 중 강력거절 및 이사추적실패가 높은 비중을 차지함에 따라, 국내외 패널조사 담당기관들은 패널유지율의 제고를 위해 다양한 노력을 경주하고 있으며, 최근 그 대안의 하나로 관심을 받는 것이 웹서베이 방식의 도입이다. 이에 본 연구의 제6장에서는 웹서베이 도입 필요성을 타 패널조사의 경험사례와 문헌연구를 통해 살펴보고, 웹서베이의 도입 가능성과 더불어 도입규모, 도입과정에 대한 논의를 진행하도록 한다. 이어 제7장에서는 혁신패널을 대상으로 웹서베이 실험연구를 진행하여 웹서베이의 모드효과(mode effect)를 분석하고, 이를 바탕으로 웹서베이의 도입 가능성 검토를 위한 실증연구를 수행한다.

「패널자료 품질개선 연구 시리즈 Ⅰ~Ⅶ」은 한국노동패널조사가 한국패널자료의 선구적 조사 및 대표 조사 자료로서의 위상을 유지케 하고, 더불어 중장기적 관점에서의 발전방안을 모색하여, 종단면 조사 자료로

서의 활용성 및 품질개선을 위한 노력의 일환으로 지금까지 이어져 오고 있다. 한국노동연구원 한국노동패널팀은 앞으로도 한국노동패널조사가 한국을 대표하는 패널조사로서, 한국의 경제 및 노동시장에 관련한 다양한 논의의 개발과 정책이슈의 개발 및 수립, 관련 연구의 지평확대에 노력을 경주할 것을 약속드린다.

제 2 장
한국노동패널조사 표본 추가 연구[1)]

제1절 연구개요

1. 연구목적

2015년까지 총 18차의 조사가 진행된 한국노동패널조사는 사회·경제적으로 급격하게 변화하는 환경 속에서 노동, 가구 경제, 관광 정책 수요를 뒷받침할 수 있는 종단자료 제공을 위한 조사로서, 그 내용은 패널을 구성하고 있는 표본가구 내 가구 구성원의 경제활동, 노동이동, 고용, 실업, 가구소득을 비롯한 각종 소득, 임금, 근로조건, 노사관계, 교육, 직업훈련 등 노동시장 전반에 걸친 주제들이다. 조사의 대상은 도시지역에 거주하는 5,000가구(1998년 기준)와 이들 가구로부터 분가한 분가가구 및 패널 노후화와 가구 대표성 보완을 위해 추가된 1,500가구(2009년) 및 가구원이다. 조사 방법은 개인면접조사(면접 타계식)로 면접원이 직접 응답

1) 본 연구는『한국노동패널 표본 추가 연구(2016)』와『한국노동패널조사(KLIPS) 표본 추가 연구 2(2017)』를 통합한 연구이다. 노동패널팀은 2016년도에 표본을 추가하기 위해 기초연구를 시작하였으나, 2015 센서스 자료 공표시기가 늦춰지면서 2016년도에 표본을 추가하지 못하고, 2017년도에 표본추가를 위한 예비조사를 실시하였다. 따라서 본고의 제1절~제4절까지는 노동패널 18차 자료를 중심으로 2016년도에 작성한 것이고, 제5절은 노동패널 19차 자료를 중심으로 2017년도에 작성한 것이다.

자를 방문하여 해당 패널을 대상으로 설문에 대한 내용을 질문하고 그 답변을 기입해 오는 방식을 이용하고 있다.

2009년에 수행된 표본 추가에도 불구하고 지속적인 패널 마모로 인하여 전체 표본 가구의 수가 매해 감소하고 있다. 또한 2009년 이후 1인 가구의 증가와 같이 지속적으로 진행되고 있는 모집단 변화를 현재 패널 표본이 온전히 대표하지 못하고 있는 것으로 판단되고 있다. 또한 패널 자료의 적극적인 정책적 활용을 위하여 요구되는 정책 대상별 통계 생산 및 대응되는 부모집단의 종단면 분석을 위해서는 충분한 수의 표본이 확보되어야 할 것이다. 이러한 노동패널에 대한 정책적 활용에 대한 요구와 모집단의 변화를 반영하기 위해서는 현행 패널에 대한 분석 및 이에 근거한 패널 추가에 대한 필요성에 대한 연구, 적정한 수준의 추가 표본 규모에 대한 연구, 그리고 실제 이를 위한 표본설계에 대한 사전연구가 필요하다.

본 연구에서는 현행 노동패널을 구성하고 있는 표본 구조를 분석하고 이에 근거하여 표본 추가를 통한 패널의 확대 및 대표성 확보에 대한 필요성을 검토한다. 이를 위하여 현행 표본의 모집단 대표성을 통계청 자료를 이용하여 분석한다. 이를 기반으로 가능한 표본 추가 방안을 제안하고 이를 위해 필요한 표본설계의 기본적인 틀을 제안하고자 한다. 이에는 적절한 표집틀의 선택, 층화 및 집락의 정의 그리고 통계적으로 부모집단 통계 작성을 위해 필요한 표본크기에 대한 내용들이 포함된다.

2. 연구내용 및 방법

한국노동패널을 구성하고 있는 표본은 크게 두 차례에 걸쳐 구성되었다. 원년인 1998년에는 도시가구를 대표할 수 있는 가구 및 가구원 표본이 구성되었으며, 이후 패널의 노후화 및 도시가구를 제외한 기타 가구의 미포함으로 발생하는 패널의 대표성 문제를 해결하기 위하여 2009년에 1,415가구와 이에 대응하는 가구원이 추가되었다. 이후에도 일정 수준의 패널 마모는 지속적으로 일어나고 있으며 이로 인한 표본크기의 감소가 통계량이나 종단면 분석 결과에 대한 통계적 신뢰도를 떨어뜨리고 있다.

본 연구에서는 이러한 문제를 극복하고 추후 지속적으로 안정적인 통계 생산과 분석을 위한 표본 추가 방안에 대한 기본적인 틀을 제공하는 것을 목적으로 다음과 같은 사항들을 검토한다.

가. 기존 연구 결과 검토

1998년과 2009년 수행된 표본추가 연구에 대한 내용 및 해외의 유사한 패널조사 사례의 표본설계 내용을 검토하고 이를 바탕으로 추가될 표본의 추출을 위한 표본설계에 반영한다. 1998년의 경우에는 노동패널 구축을 위한 표본설계가 별도로 이루어지지 않았으며 이를 보완하기 위한 가중치 작업이 추후에 이루어졌으나, 2009년 표본추가를 위해서는 별도의 표본설계 연구가 이루어졌고 원패널과 2009년 신규 패널의 합집합으로 정의되는 통합패널의 모집단 대표성이 일정 수준 이상 이 과정을 통해 확보되었다. 본 연구에서는 2009년 표본설계 내용을 검토하고 이를 반영한 표본설계(안)을 제시하려 한다. 이는 생산되는 통계와 분석 결과에 큰 영향을 미치는 표본설계가 극단적으로 변경될 경우에 통계 및 분석 결과의 안정성에 영향을 미치기 때문이다. 또한 노동패널과 유사한 해외의 패널조사 사례, 예를 들어, 미국의 PSID(Panel Study of Income Dynamics), 호주의 HILDA(Household, Income and Labor Dynamics in Australia), 영국의 BHPS(British Household Panel Study) 그리고 독일의 G-SOEP (German Socio-Economic Panel) 등을 살펴보고자 한다. 특히, 이러한 패널들을 구성하기 위하여 사용된 표본설계 방안과 원리를 검토하여 새롭게 추가될 패널의 추출을 위한 표본설계에 반영한다.

나. 현행 패널의 구성 분석

신규 패널 추가의 목적은 언급한 바와 같이 분석의 정밀화와 표본의 모집단 대표성 확보에 있다. 이를 위해서는 기본적으로 현행 패널 구성에 대한 검토가 선행되어야 한다. 본 연구에서는 현재 사용되고 있는 패널의 표본구성에 대한 다양한 분석을 실시하여 신규 패널 구성을 위한 표본 추

가의 필요성을 검토한다. 이를 위해 기본적인 가구 및 가구원의 특성들에 대한 분포를 대응되는 시점의 모집단 분포와 비교하고, 또한 기 산출된 가중치 분석을 통해서 패널 마모로 인하여 발생한 무응답이 추정에 미치는 영향력을 판단하여 표본의 추가를 위해 필요한 기본적인 정보를 제공하려 한다.

다. 신규 패널 구축을 위한 표본설계 방안의 틀 제안

본 연구에서는 기존 연구 및 현행 패널 구조에 대한 분석을 바탕으로 신규로 구축하려 하는 패널을 위한 표본설계 방안의 기본적인 틀을 제안한다. 특히 2015년 수행된 등록센서스 기반 조사구 표집틀이 확보되기 전인 현시점에서 사용 가능한 정보를 통해서 가능한 한 구체적인 표본설계 방안을 제시하고 이를 근거로 실제 표집틀이 주어졌을 때 표본추출이 가능할 수 있도록 한다.

제2절 현행 노동패널 표본의 모집단 대표성 분석

1. 개 요

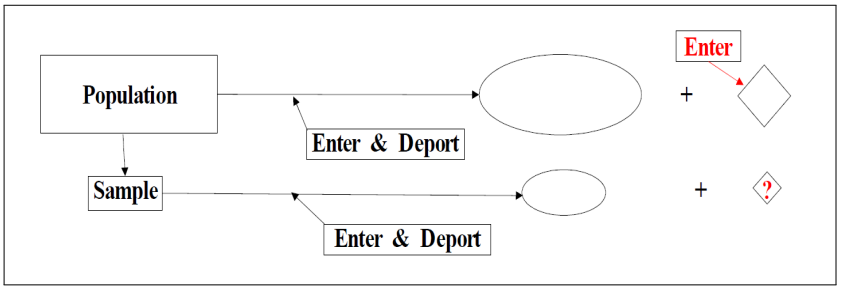
패널조사를 포함한 대부분의 조사 자료 분석 결과는 조사된 표본 및 사용하는 통계 분석 방안에 의하여 그 효율성이 결정된다. 이러한 관점에서 표본설계의 중요한 목적 중 하나는 구현된 표본을 이용한 분석 결과의 통계적 타당성을 확보함과 동시에 효율성을 높이는 것이다. 통계적 타당성과 효율성 측면에서 표본이 갖추어야 할 중요한 성질로는 흔히 모집단 대표성(representativeness)이 고려된다. 보조정보(auxiliary information)를 활용한 확률 표본설계의 목적은 대부분 임의표본(random sample)의 조건을 만족함과 동시에 모집단 대표성을 확보하는 것에 있다. 이러한 대표적인 표본설계로는 층화추출법 그리고 표집틀의 정렬 후 수행하는 계

통추출 등이 있다.

모집단 대표성이란 기본적으로 표본이 모집단의 모방(mimic)이 되는 것을 의미한다. 즉 주어진 변수 기준으로 모집단의 분포와 표본의 분포가 유사할 때 우리는 이러한 표본을 대표성을 갖춘 표본이라 정의하게 된다. 만약 조사를 통해 분석하고자 하는 변수 기준으로 대표성을 갖춘 표본을 추출할 수 있다면, 연구자는 오차가 거의 없는 모집단에 대한 추론을 할 수 있다. 그러나 대부분의 서베이에서 관심의 대상이 되는 변수 혹은 문항은 그 수가 매우 많고 또한 이에 대한 모집단 분포 자체가 추론의 대상이기 때문에 설문 문항 변수를 기준으로 표본의 모집단 대표성을 평가할 수는 없다. 따라서 많은 경우 표본의 모집단 대표성을 평가하기 위해서 인구사회학적 특성 변수를 고려하게 된다. 보다 전문적인 조사의 경우에는 조사와 관련성이 높은 변수를 추가적으로 고려할 수 있으나 이를 위해서는 해당 변수에 대한 모집단 정보가 주어져야 한다. 개인 단위 조사에서는 흔히 거주 지역, 성별 그리고 연령의 정보가 대표성 평가를 위해 흔히 고려되며, 가구 조사의 경우에는 지역, 가구 형태(일반/아파트, 1인 가구/2인 가구 등) 등이 고려된다. 이는 적어도 이러한 변수들이 설문항목과 정도의 차이는 있으나 유의한 연관을 가지고 있고 따라서 이러한 인구사회학적 특성 변수 측면에서의 대표성 확보가 조사 결과를 이용한 분석의 효율성을 높일 수 있다고 가정하기 때문이다.

언급한 대표성의 개념은 1회성 혹은 횡단면 분석을 위한 표본설계 및 이를 통해 구현되는 표본에 적용할 수 있다. 통상적으로 1회의 표본추출을 통해 구성된 표본을 반복적으로 조사하는 패널조사의 경우에는 그 특성상 시간의 개념을 포함한 대표성 역시 고려하여야 한다. 패널조사의 구조는 다음의 [그림 2-1]과 같이 단순화할 수 있다. 종단면 분석을 위해 구성되는 패널의 경우에 종단면 측면에서의 대표성의 개념은 시간의 경과에 따른 모집단 변화(population change)의 표본 반영 여부로 이해할 수 있다. [그림 2-1]에서와 같이 패널 구성 시에 직사각형의 모습을 갖춘 모집단은 시간이 경과함에 따라 모집단으로부터의 이탈 혹은 새로운 개체의 유입으로 그 형태가 바뀌게 된다. 이탈의 예로는 사망과 해외 이주 등을 생각할 수 있고, 유입으로는 입양 혹은 출생 등으로 인한 경우를 고

〔그림 2-1〕 패널조사의 구조



려할 수 있을 것이다. 이러한 모집단의 변화가 일정 수준 표본에서도 동일하게 이루어지고 있다고 가정할 수 있을 때 종단면 측면에서의 패널이 대표성을 갖추었다고 표현할 수 있을 것이다. 그러나 횡단면 측면의 대표성의 논의 시 언급한 바와 같이 종단면 측면에서의 대표성을 평가하기 위해서는 모집단의 변화를 관측할 수 있는 보조정보가 필요하고 이러한 정보는 대부분 인구사회학적 항목으로 제한되고 있다.

본장에서는 언급한 패널의 모집단 대표성을 평가하기 위하여 기본적인 가구원의 인구사회학적 정보 및 가구 정보를 고려하였다. 이를 위해서 기본적으로 구축시기(1998년/2009년) 그리고 분석목적(횡단면/종단면)의 총 네 가지 범주에서 시간의 경과에 따른 유효 표본크기를 살펴보고 18차 자료를 조사 해당 연도인 2015년 통계청 추계 자료와 비교하여 패널 대표성 및 패널 추가의 필요성을 평가한다.

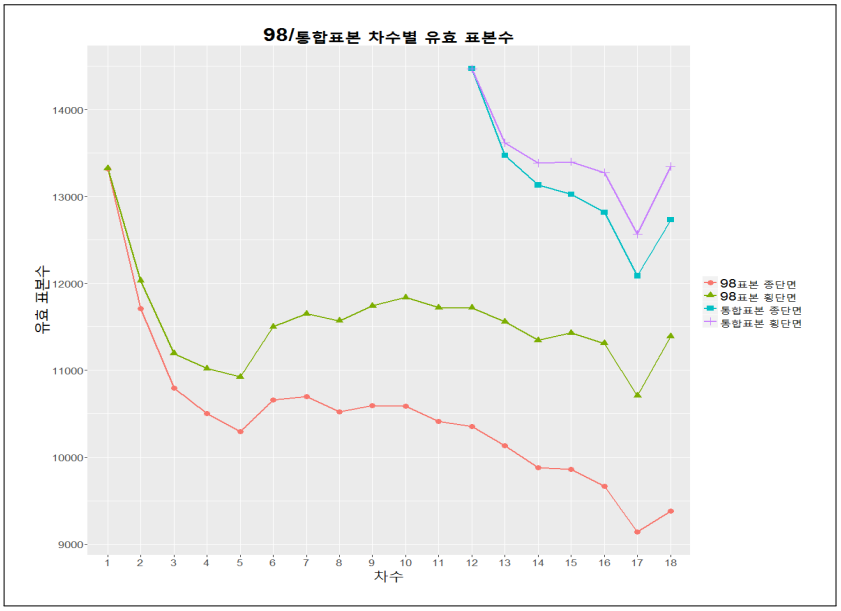
2. 현행 노동패널의 표본 가구원 분석

패널을 이용한 분석 방안은 크게 종단면 분석과 횡단면 분석으로 구분할 수 있다. 횡단면 분석(cross-sectional analysis)이란 조사시점의 유한 모집단에 대한 추론을 위한 분석을 의미하고, 종단면 분석(longitudinal analysis)이란 조사시점까지의 패널의 변동에 대한 동적(dynamic) 분석을 의미한다. 노동패널 자료의 분석 역시 두 측면에서 이루어지며 매해 횡단면 가중치와 종단면 가중치가 원패널 및 통합패널의 분석을 위해 산출된다. 즉 2009년 이후 매해 네 종류의 가중치가 산출되어 제공되고 있다.

[그림 2-2]와 <표 2-1>은 네 가지 종류의 가중치가 각 차수별로 실제 부여되는 대상의 수를 나타내는 유효 가구원 표본수를 나타내고 있다. 종단면 가중치는 조사 원년부터 해당 시점까지 조사된 가구원에게 부여되며, 횡단면 가중치는 해당 시점에 조사된 모든 가구원에게 부여된다. 따라서 횡단면 가중치가 부여되는 표본 가구원수는 항상 종단면 가중치가 부여되는 가구원수보다 크게 나타나고 있다. 종단면 가중치가 부여되는 가구원수가 단조 감소(monotone decreasing) 형태가 아닌 것은 각 조사시점에 원가구로부터 분가한 가구나 원패널 가구에 속한 가구원이면서 조사시점에 연령이나 기타 기준으로 패널 조건을 만족하여 포함되는 경우가 발생하기 때문이다. 원 시점인 1998년 혹은 2009년의 표본수와 비교할 때 18차에 해당하는 유효 표본수는 매우 적음을 알 수 있다. 특별히 종단면 기준으로 볼 때 1998년을 원 시점으로 간주하여 종단면 분석을 실시할 경우에 유효 가구원 표본수는 매우 적음을 확인할 수 있다.

[그림 2-3]은 통계청 제공 추계자료를 이용하여 산출한 지역, 성과 연령의 결합분포와 18차(2015년)에 조사되고 횡단면 가중치가 부여된 가구

〔그림 2-2〕 원패널 및 통합패널의 유효 표본 가구원수



〈표 2-1〉 원패널 및 통합패널의 유효 표본 가구원수

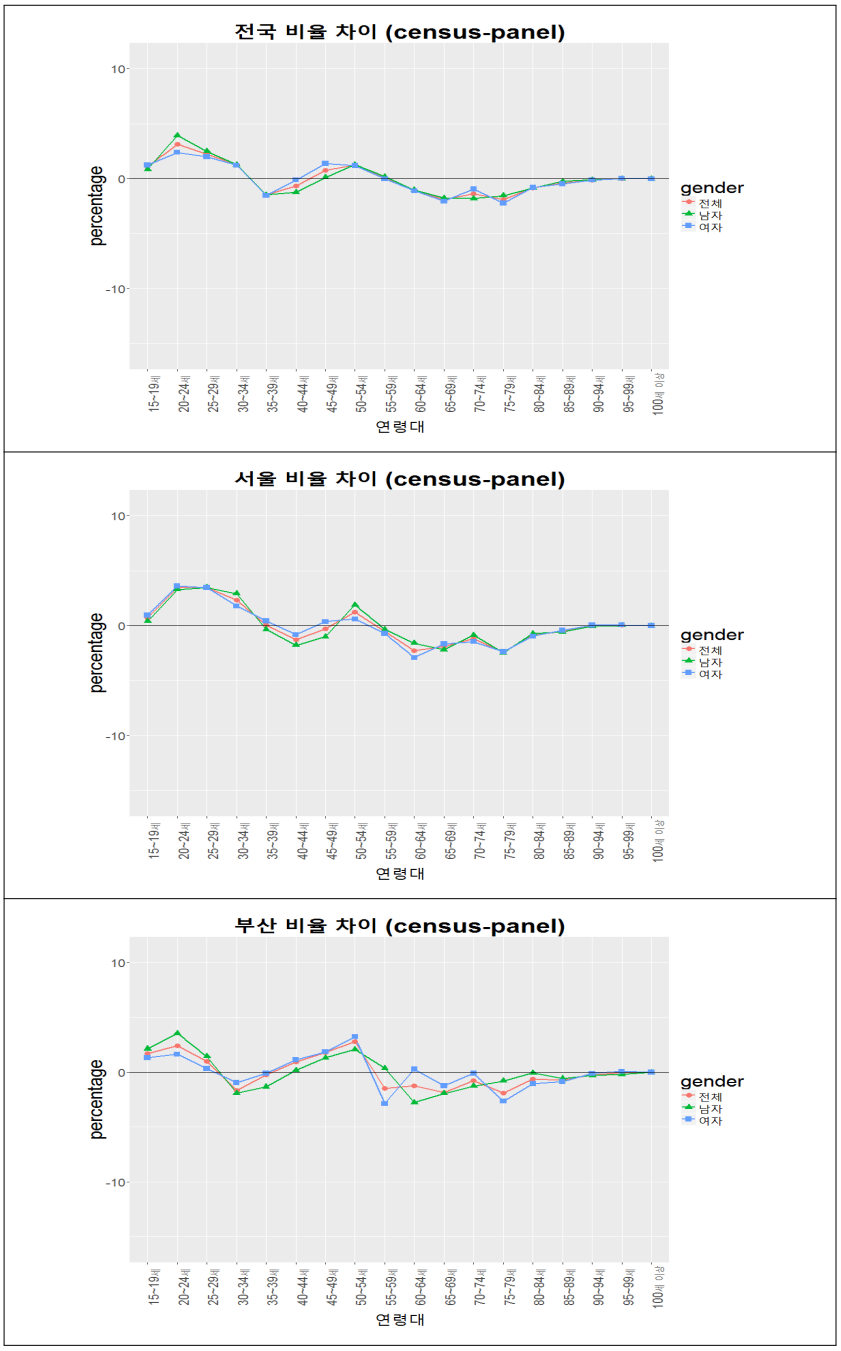
	98표본		통합표본	
	종단면	횡단면	종단면	횡단면
1	13,321	13,321		
2	11,708	12,031		
3	10,797	11,194		
4	10,500	11,021		
5	10,296	10,924		
6	10,659	11,501		
7	10,696	11,651		
8	10,522	11,570		
9	10,591	11,741		
10	10,590	11,840		
11	10,411	11,720		
12	10,356	11,719	14,470	14,470
13	10,132	11,560	13,472	13,618
14	9,878	11,347	13,132	13,385
15	9,861	11,428	13,024	13,400
16	9,669	11,310	12,818	13,276
17	9,142	10,707	12,086	12,566
18	9,379	11,391	12,731	13,345

원 패널의 지역, 성과 연령의 결합분포의 차이를 나타내고 있다. 즉 그림의 각 점들은 해당 성과 연령 그룹에 속하는 비율의 차이로 다음과 같이 표현할 수 있다.

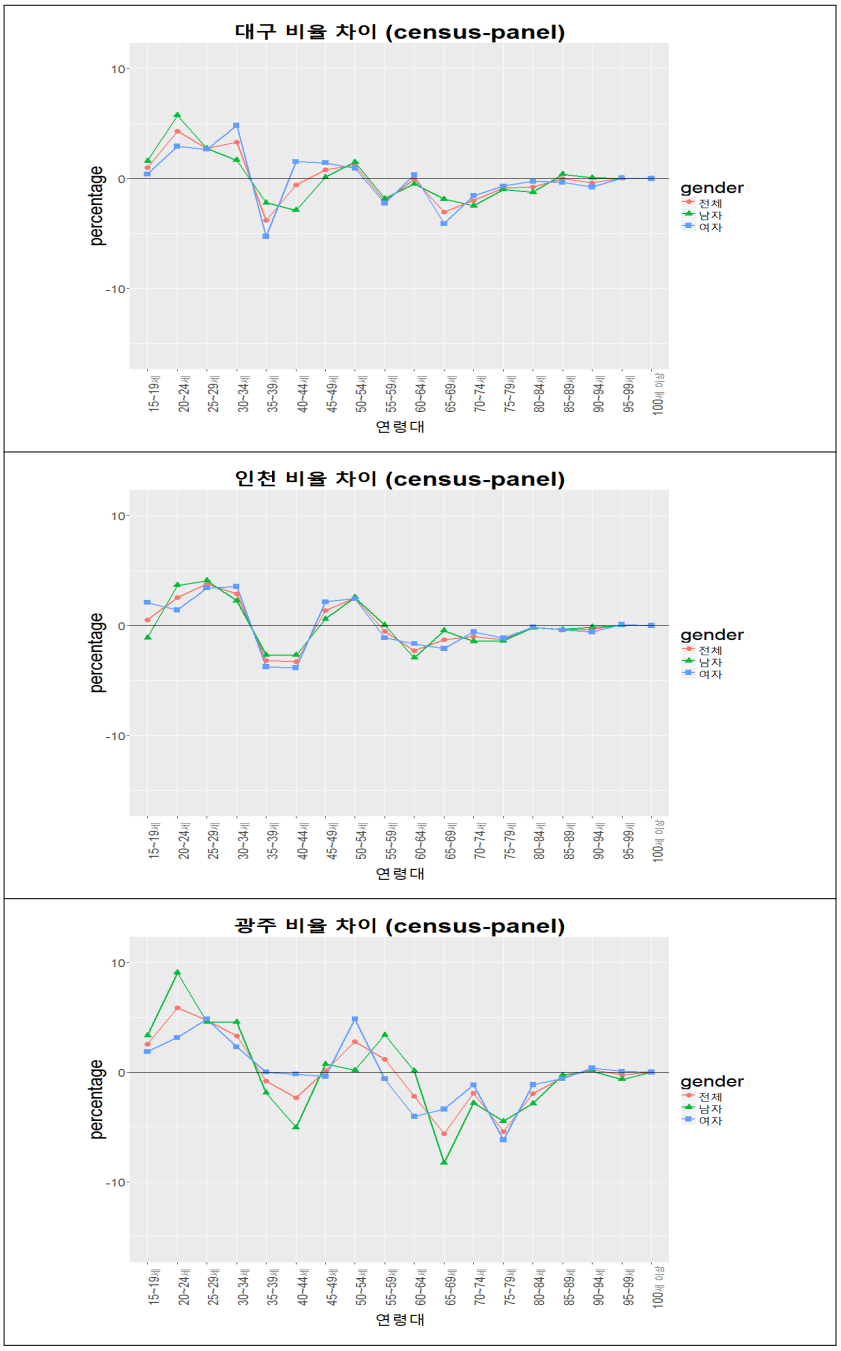
$$p_{\text{추계}} - p_{\text{노동패널}}$$

신규 행정도시로서 특별한 성격을 갖는 세종시를 제외하고는 대체적으로 모든 지역에서 동일한 패턴을 나타내고 있다. 다만 광주와 제주의 경우는 다른 지역에 비하여 연령 그룹 간 추계와의 차이의 변동이 크게 나타나고 있음을 알 수 있다. 인구추계결과와 비교할 때 전반적으로 저연령대의 패널 비율이 유의하게 낮고 50세 이상의 고연령대 패널 비율이

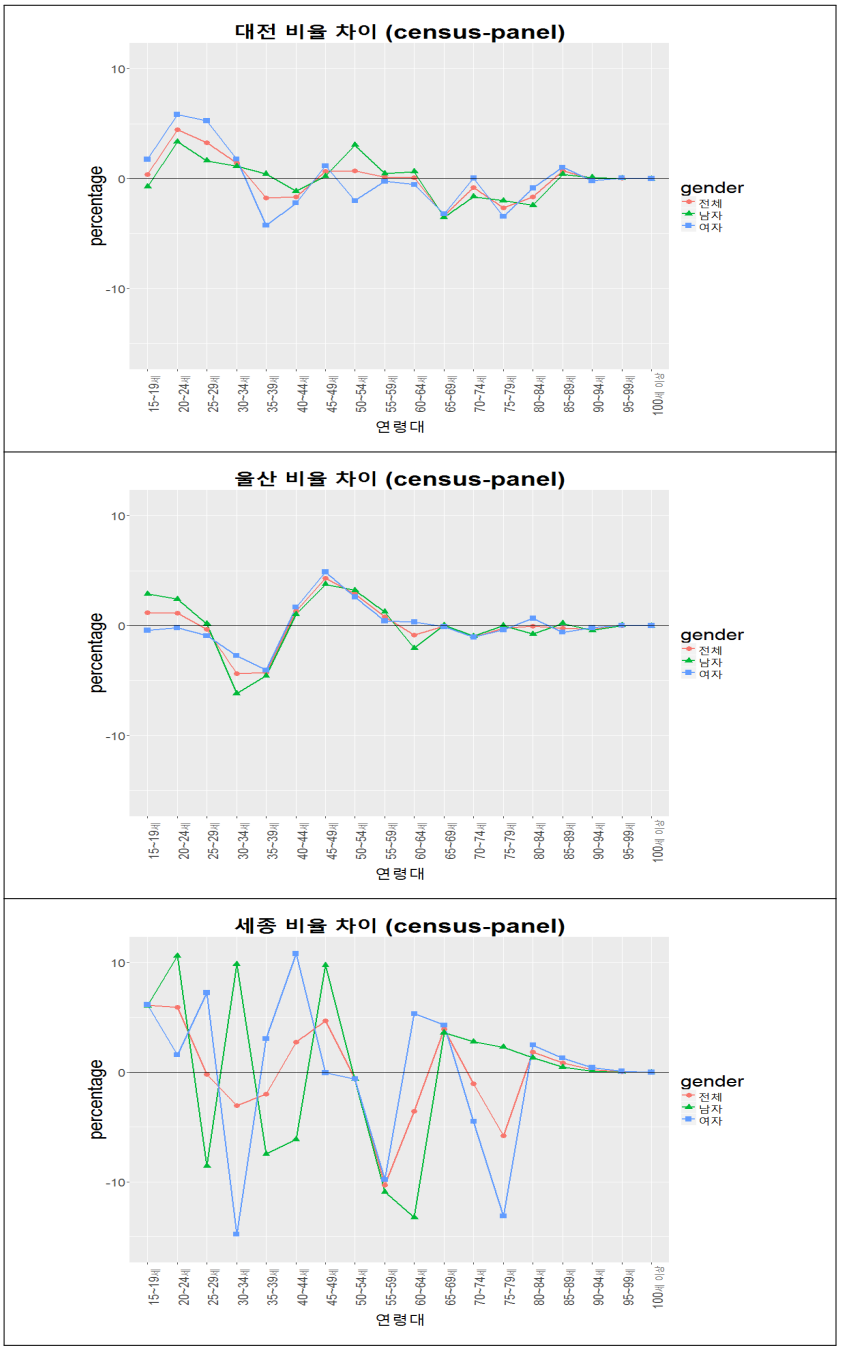
〔그림 2-3〕 성*연령에 대한 추계자료와 패널자료의 비교



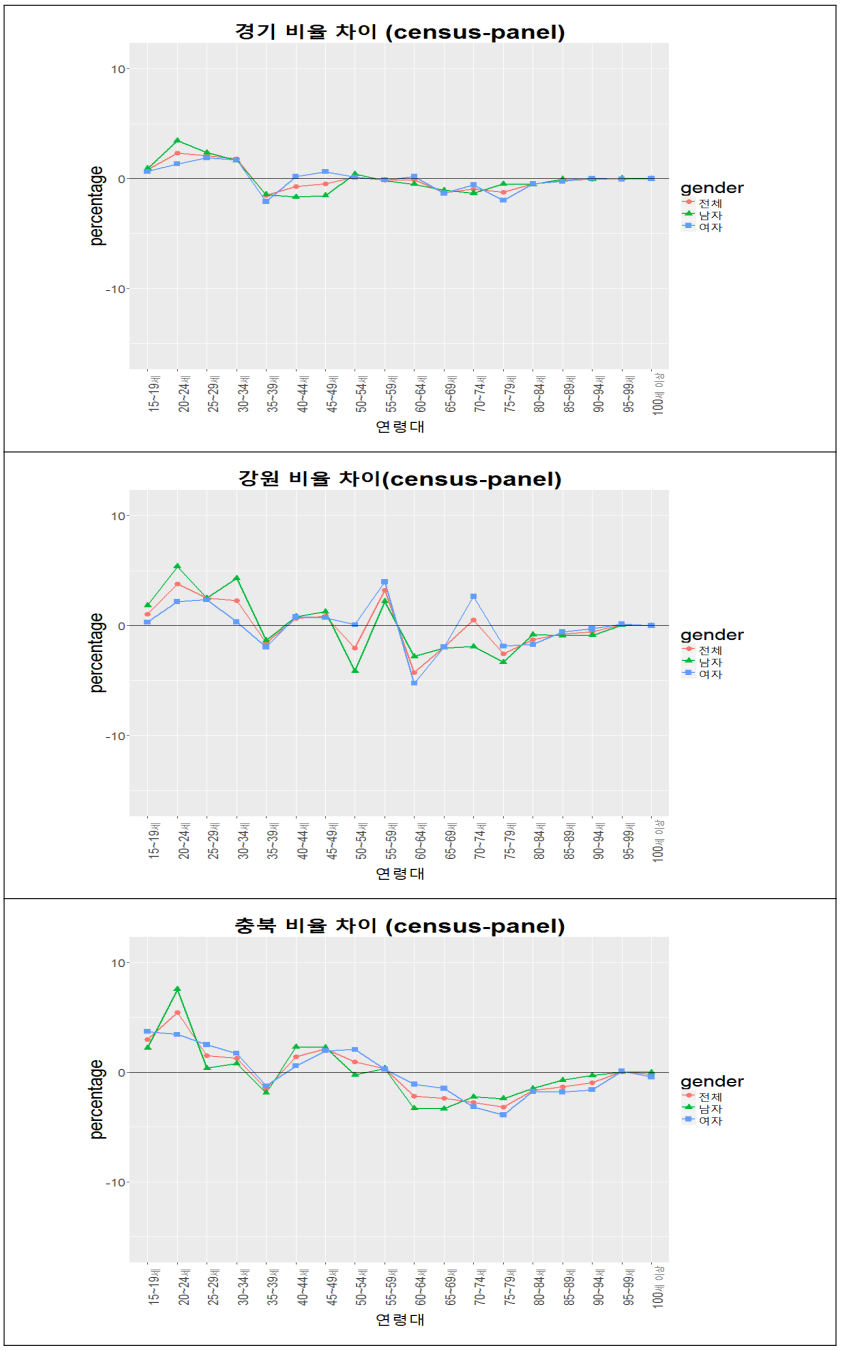
〔그림 2-3〕의 계속


대구 비율 차이 (census-panel)
인천 비율 차이 (census-panel)
광주 비율 차이 (census-panel)
percentage
10
0
-10
15~19세
20~24세
25~29세
30~34세
35~39세
40~44세
45~49세
50~54세
55~59세
60~64세
65~69세
70~74세
75~79세
80~84세
85~89세
90~94세
95~99세
100세 이상
연령대
gender
전체
남자
여자

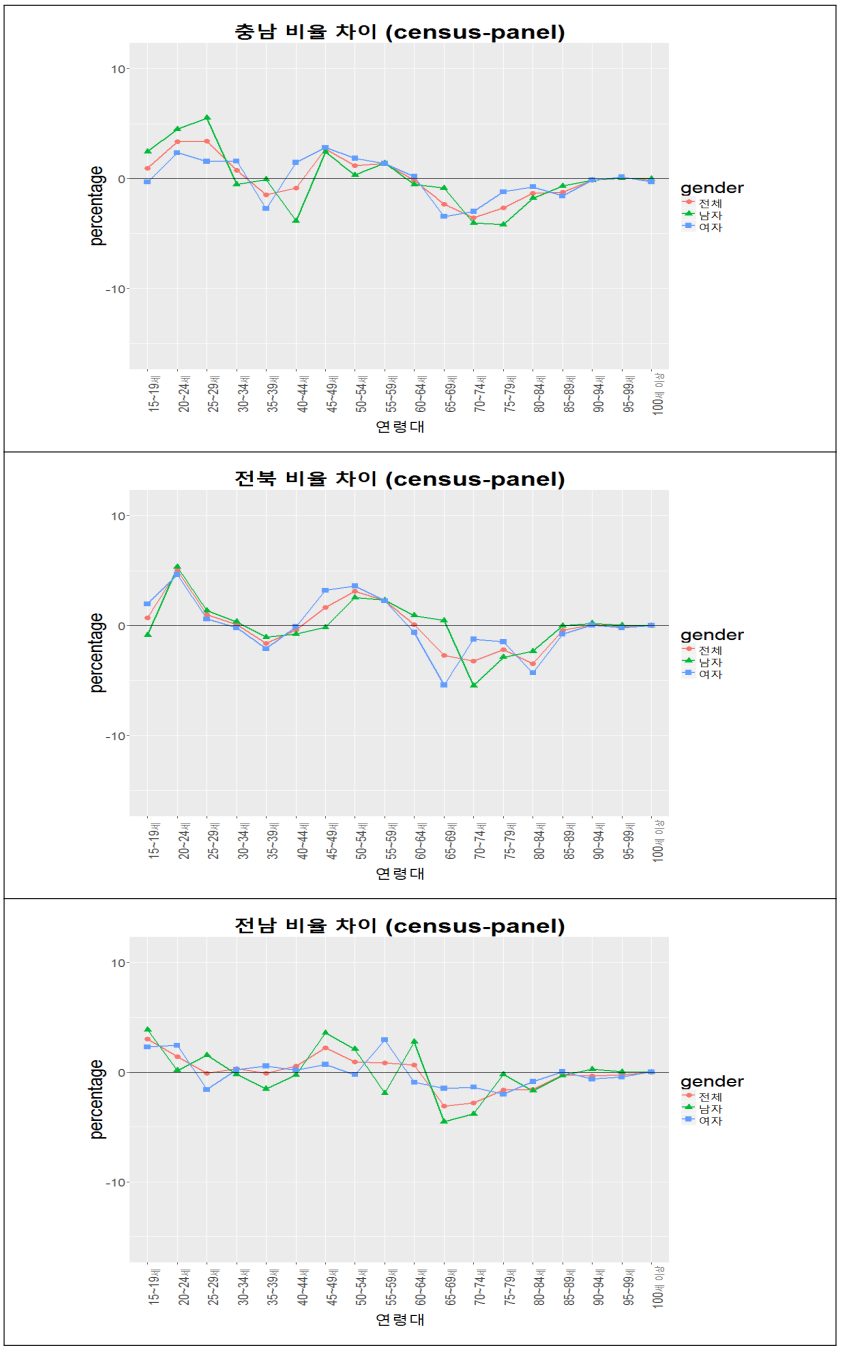
〔그림 2-3〕의 계속


대전 비율 차이 (census-panel)
울산 비율 차이 (census-panel)
세종 비율 차이 (census-panel)
percentage
10
0
-10
15-19세
20-24세
25-29세
30-34세
35-39세
40-44세
45-49세
50-54세
55-59세
60-64세
65-69세
70-74세
75-79세
80-84세
85-89세
90-94세
95-99세
100세 이상
연령대
gender
전체
남자
여자

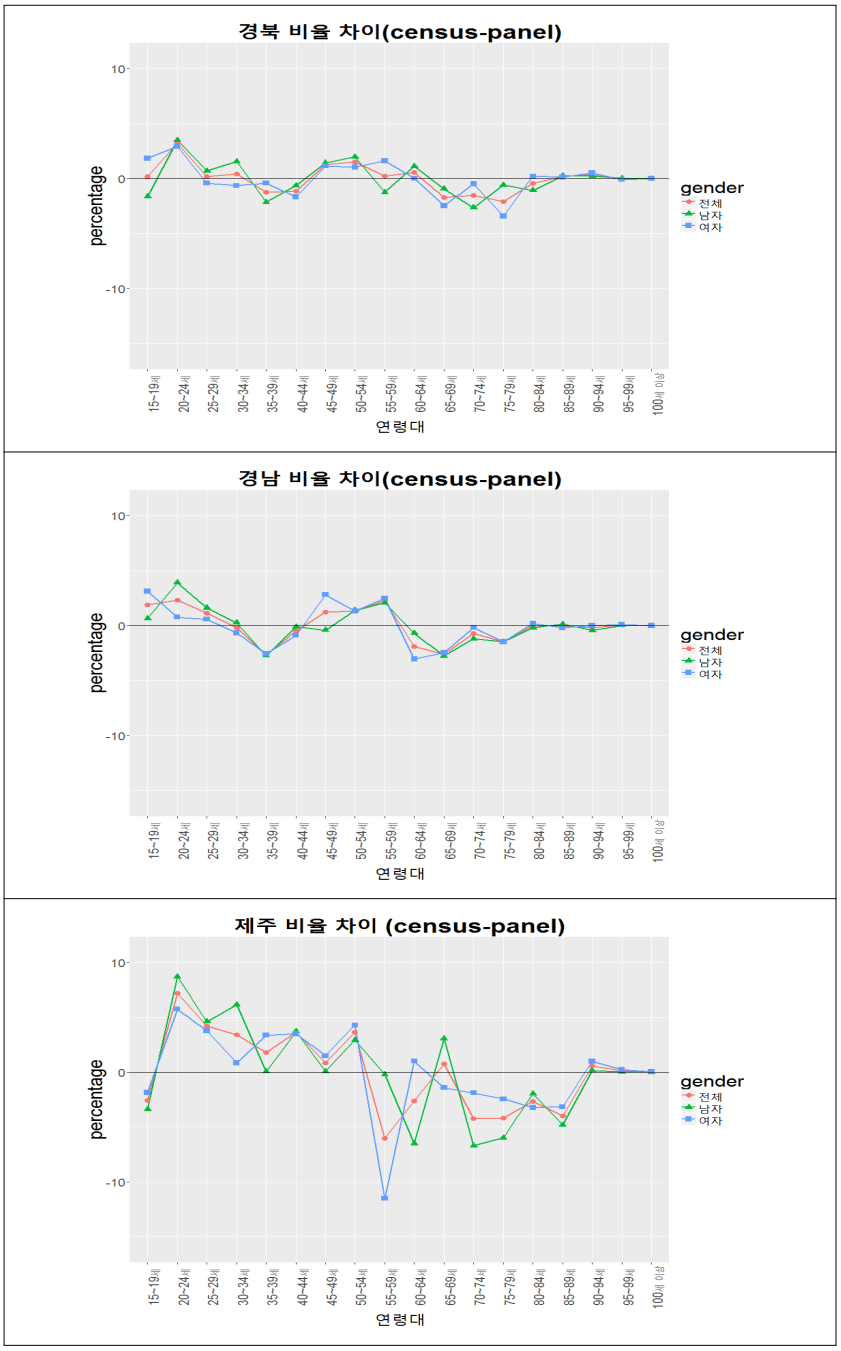
〔그림 2-3〕의 계속


경기 비율 차이 (census-panel)
강원 비율 차이(census-panel)
충북 비율 차이 (census-panel)
percentage
10
0
-10
15-19세
20-24세
25-29세
30-34세
35-39세
40-44세
45-49세
50-54세
55-59세
60-64세
65-69세
70-74세
75-79세
80-84세
85-89세
90-94세
95-99세
100세 이상
연령대
gender
전체
남자
여자

〔그림 2-3〕의 계속


충남 비율 차이 (census-panel)
전북 비율 차이 (census-panel)
전남 비율 차이 (census-panel)
percentage
10
0
-10
15-19세
20-24세
25-29세
30-34세
35-39세
40-44세
45-49세
50-54세
55-59세
60-64세
65-69세
70-74세
75-79세
80-84세
85-89세
90-94세
95-99세
100세 이상
연령대
gender
전체
남자
여자

〔그림 2-3〕의 계속


경북 비율 차이(census-panel)
percentage
10
0
-10
15~19세
20~24세
25~29세
30~34세
35~39세
40~44세
45~49세
50~54세
55~59세
60~64세
65~69세
70~74세
75~79세
80~84세
85~89세
90~94세
95~99세
100세 이상
연령대
gender
전체
남자
여자
경남 비율 차이(census-panel)
percentage
10
0
-10
15~19세
20~24세
25~29세
30~34세
35~39세
40~44세
45~49세
50~54세
55~59세
60~64세
65~69세
70~74세
75~79세
80~84세
85~89세
90~94세
95~99세
100세 이상
연령대
gender
전체
남자
여자
제주 비율 차이 (census-panel)
percentage
10
0
-10
15~19세
20~24세
25~29세
30~34세
35~39세
40~44세
45~49세
50~54세
55~59세
60~64세
65~69세
70~74세
75~79세
80~84세
85~89세
90~94세
95~99세
100세 이상
연령대
gender
전체
남자
여자

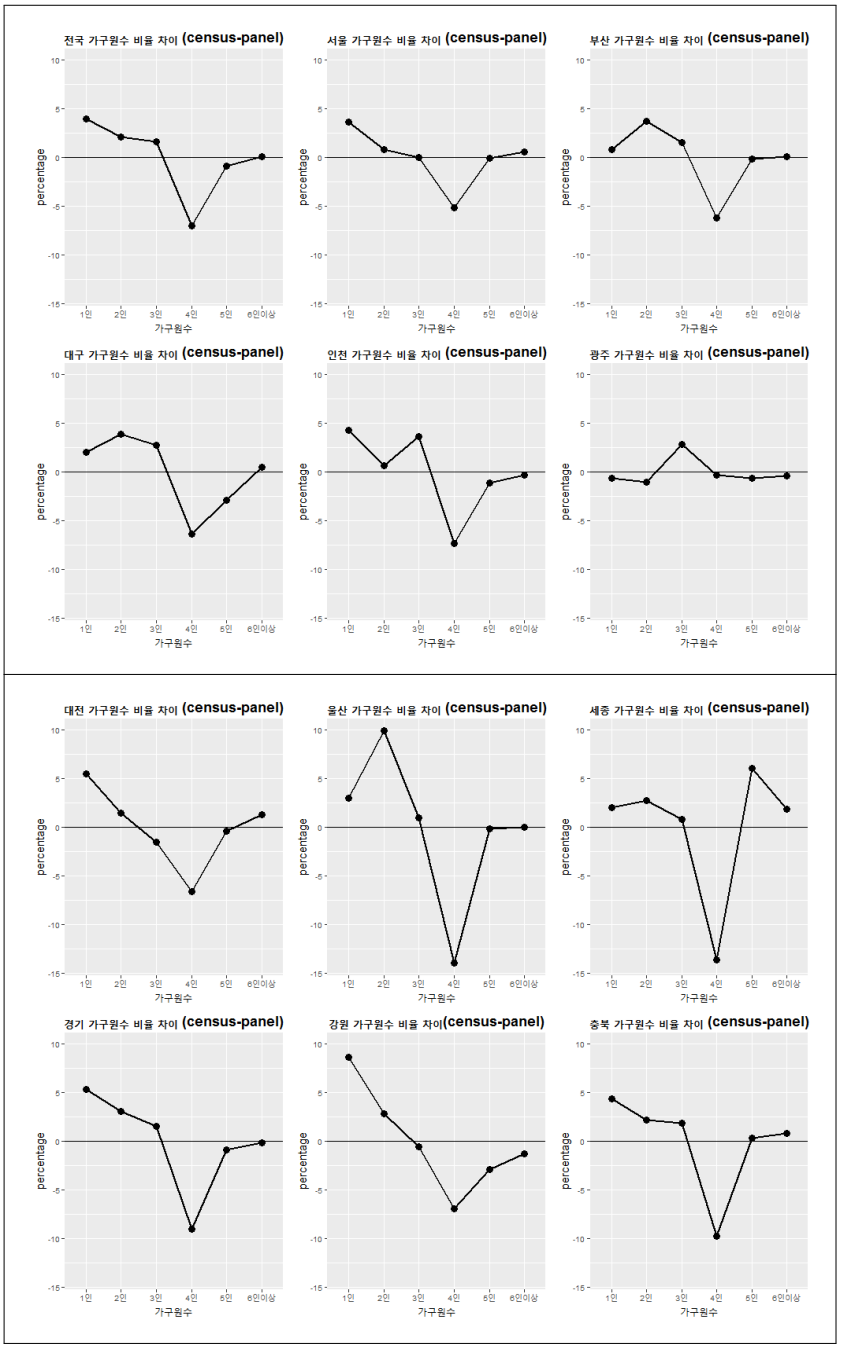
높음을 알 수 있다. 또한 30대의 비율 역시 지역 간 변동은 있으나 패널의 경우가 더 높게 나타나고 있음을 확인할 수 있다. 전체적으로 두 비율의 차이는 5% 수준 내외로 나타나고 있으나 경우에 따라서는 10% 이상의 차이를 보이는 범주도 있음을 알 수 있다.

조사자료를 이용한 추정량의 분산은 크게 관심변수의 변동과 생성되는 가중치의 변동에 의존하게 된다. 만약 두 변동 간에 음의 상관관계가 성립한다면 추정량의 분산을 최소화할 수 있으나 이는 실제로는 가능하지 않다. 대신 가능한 범위에서 가중치의 변동을 최소화하는 것이 분석의 효율성을 높이기 위해서 바람직하다. 가중치의 변동을 최소화하기 위해서는 표본의 모집단 대표성을 확보하여 자료 수집 이후에 무응답 조정이나 캘리브레이션 과정에서 발생하는 추가적인 변동을 줄여야 한다. 특히 노동패널조사의 경우에 캘리브레이션을 위하여 기본적으로 언급한 가구원의 인구사회학적 변수가 사용되기 때문에 이러한 변수 기준 표본의 모집단 대표성을 확보하는 것이 요구된다. 더불어 대부분의 종단면 분석을 위해서는 혼합선형모형과 같은 특정한 통계 모형이 가정되기 때문에 이러한 가정을 왜곡(torture)하지 않는 표본 추출법이 사용되는 것이 바람직하다. 이를 위해서는 각 개인이 부여받는 가중치의 개인 간 변동이 최소화되는 것이 바람직하며 이를 위해서도 가중치 변동을 최소화할 수 있도록 패널의 구조를 구성하는 것이 바람직하다.

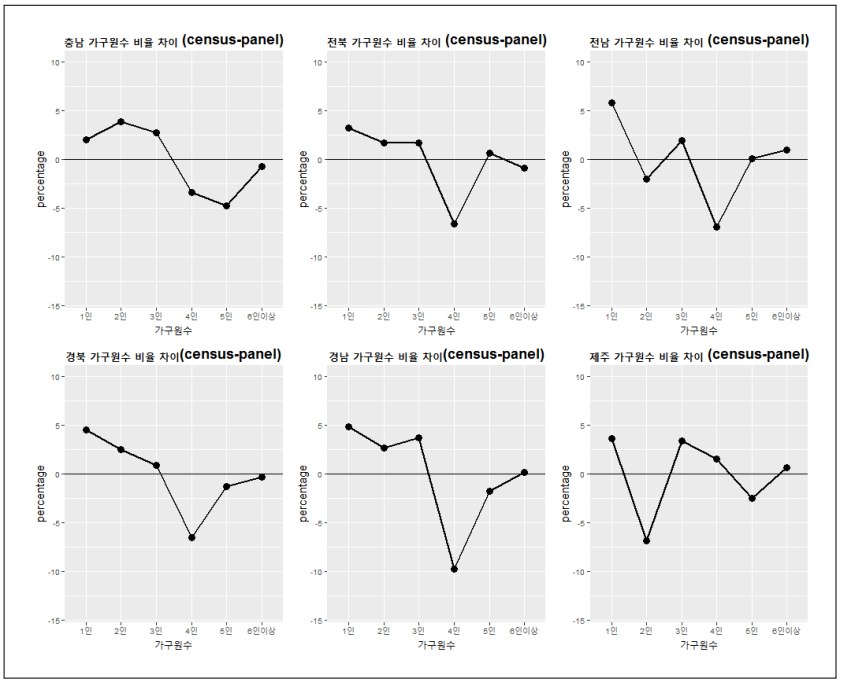
3. 현행 노동패널의 표본가구 분석

[그림 2-4]는 통계청의 가구추계결과와 18차 노동패널의 가구 분포의 차이를 나타내고 있다. 가구 분포를 나타내기 위해서 가구 구성원수 기준으로 1인 가구, 2인 가구, 3인 가구, 4인 가구, 5인 가구 그리고 6인 이상 가구의 범주를 사용하였다. 예상할 수 있듯이 1인 혹은 2인 가구의 비율이 노동패널에서 상당히 낮게 나타나고 있다. 이는 모집단에서의 1인 혹은 2인 가구의 증가와 기존 패널의 가구 특성이 나타난 결과로 판단된다. 다만 지역적으로 다소 차이가 나타나고 있다. 특히 인구 분포에서와 마찬가지로 광주와 제주 지역의 경향이 다른 지역과 다르게 나타나고 있어

〔그림 2-4〕 가구 유형에 대한 추계자료와 패널자료의 비교



〔그림 2-4〕의 계속

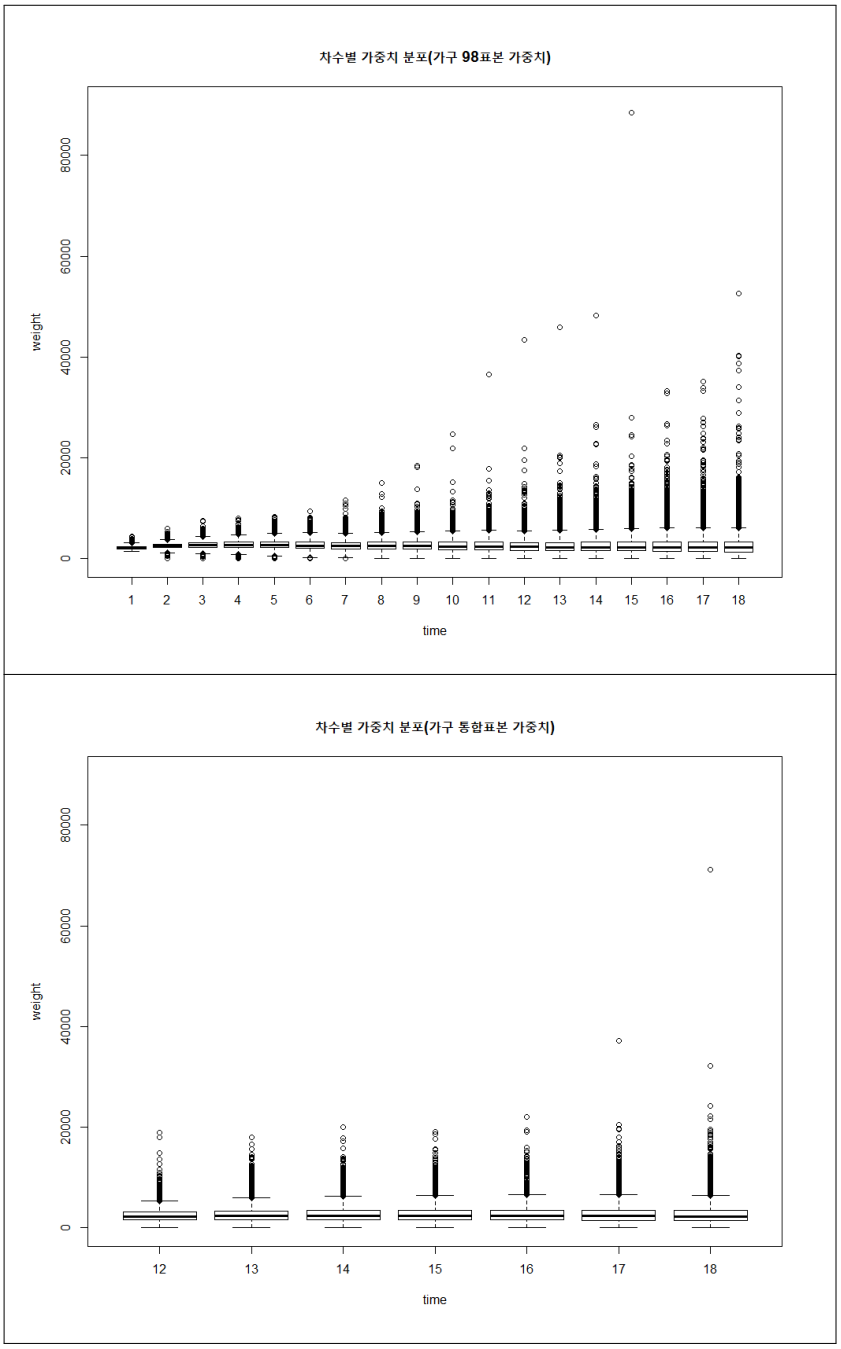


추후 패널 추가 시 이를 고려하는 것이 요구된다. 전반적으로 새로운 패널의 구축 시 현재 모집단의 가구 분포를 반영하여 1인 혹은 2인 가구를 패널에 적극적으로 포함하여야 할 것으로 판단된다. 특히 노동시장에 새롭게 진입했거나 또는 곧 진입 예정인 개인이 속한 가구가 주로 1인 혹은 2인 가구임을 고려할 때 이를 반영한 표본 추가가 이루어지는 것이 바람직할 것이다.

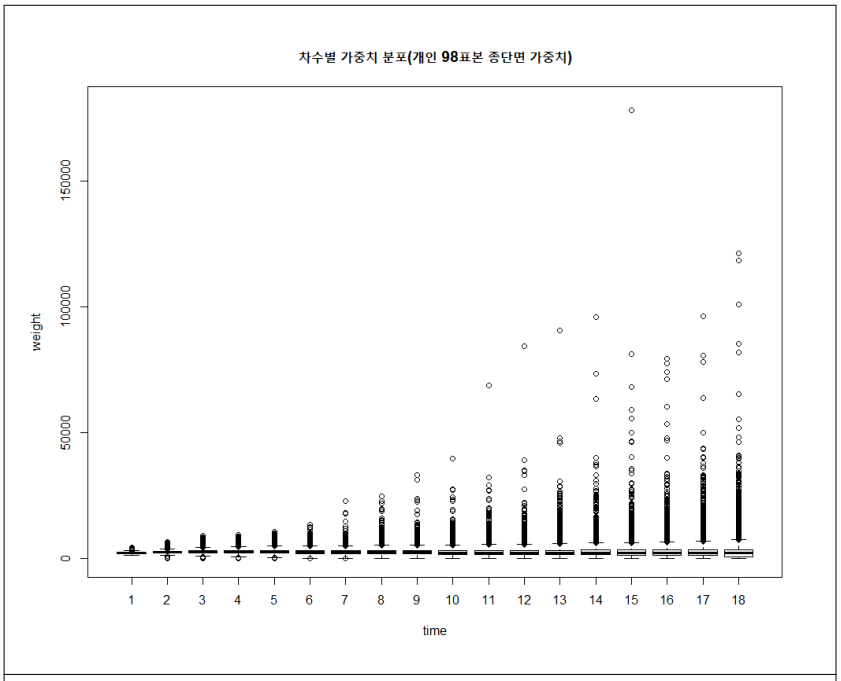
4. 현행 노동패널의 가중치 분석

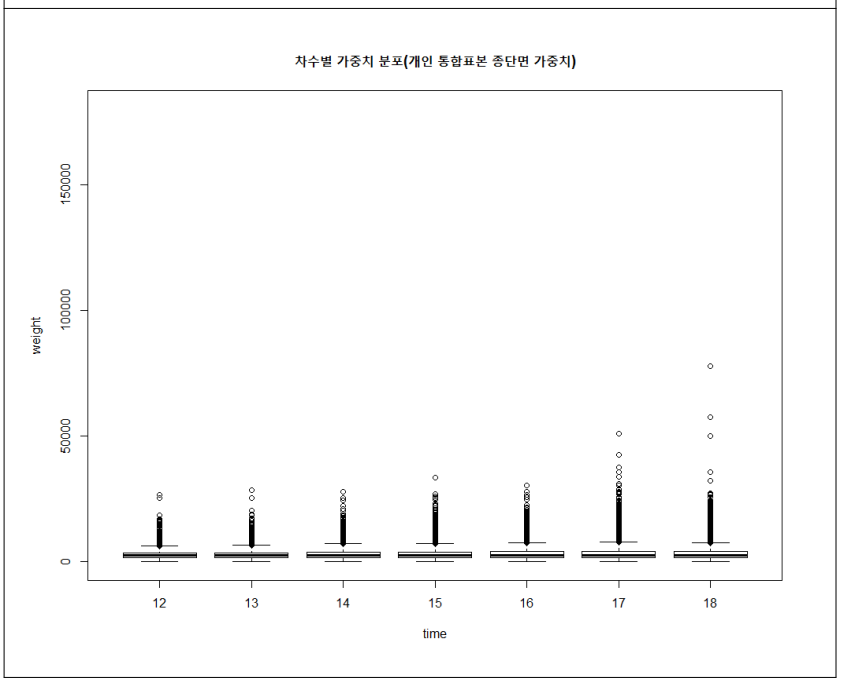
[그림 2-5]는 1차부터 현재까지 원패널과 통합패널의 횡단면 및 종단면 분석을 위해 작성된 가중치의 병렬 상자그림(side-by-side boxplot)을 나타내고 있다. 2항에서 살펴본 바와 같이 시간의 경과에 따라 유효 표본의 수가 감소하고 따라서 각 시점에서 실제 응답한 개인이 대표해야 하는

〔그림 2-5〕 가중치 분포

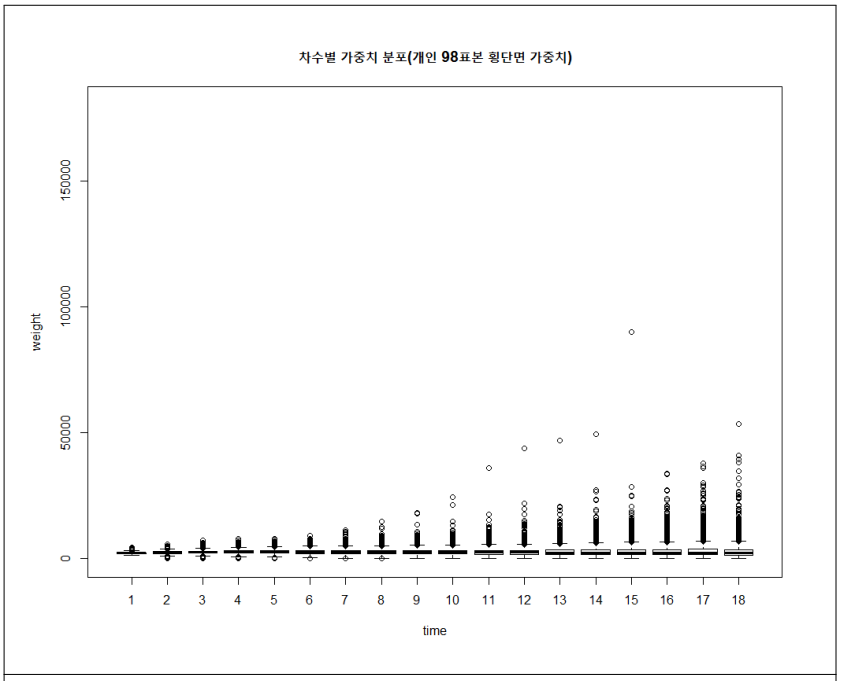

차수별 가중치 분포(가구 98표본 가중치)
weight
0
20000
40000
60000
80000
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
time
차수별 가중치 분포(가구 통합표본 가중치)
weight
0
20000
40000
60000
80000
12 13 14 15 16 17 18
time

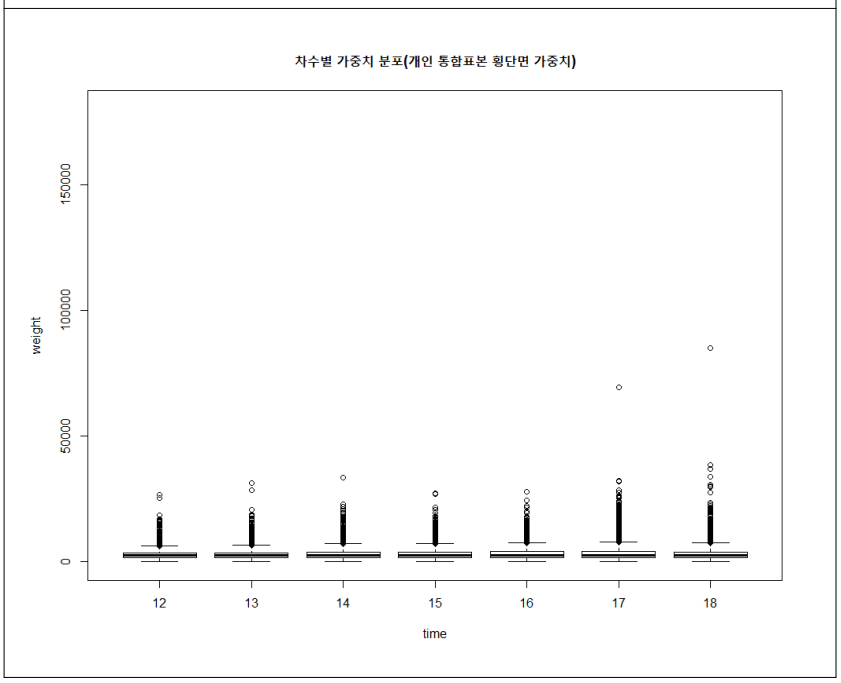
〔그림 2-5〕의 계속


차수별 가중치 분포(개인 98표본 종단면 가중치)
weight
0
50000
100000
150000
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
time


차수별 가중치 분포(개인 통합표본 종단면 가중치)
weight
0
50000
100000
150000
12 13 14 15 16 17 18
time

〔그림 2-5〕의 계속


차수별 가중치 분포(개인 98표본 횡단면 가중치)
weight
0
50000
100000
150000
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
time


차수별 가중치 분포(개인 통합표본 횡단면 가중치)
weight
0
50000
100000
150000
12 13 14 15 16 17 18
time

개인의 수는 증가하게 되고 이로 인하여 각 가중치는 증가하고 있음을 알 수 있다. 특히 종단면 가중치의 경우 그 가중치의 증가 경향이 더욱 명확하게 드러나고 있으며 패널 추가가 이루어진 통합패널보다 원패널의 가중치 증가가 더욱 뚜렷하게 나타나고 있다. 한 가지 중요한 현상은 가중치 중 특별히 매우 큰 값을 갖는 이상치가 나타나는 것으로 이는 패널 이탈이 특정한 인구사회학적 특성을 가진 개인에게 지속적으로 발생하여 무응답 조정 과정이나 벤치마킹 과정에서 이들에게 부여된 가중치가 지속적으로 증가하고 있음을 나타내고 있다. 이러한 경향은 실제 유효 표본수를 살펴본 2항의 결과에서 확인할 수 있다.

제3절 사례연구 : 해외 패널조사에서의 표본 보완 및 개편 사례

장기간 추적하여 조사를 수행하는 패널조사는 시간의 흐름에 따라 패널의 탈락, 거부 및 추적 실패 등의 이유로 인한 패널 마모가 발생되고, 모집단 커버리지의 개선, 특정 계층이나 그룹의 모집단 변화 반영 및 표본크기의 확대 등으로 인하여 표본을 조정(refresh)하거나 추가하는 상황이 발생하게 된다. 본절에서는 국외 패널조사 중 표본을 조정하거나 추가하는 등의 표본 보완 및 개편 사례를 요약해 정리해 보고자 한다.

1. 미국의 PSID

미국의 PSID(Panel Survey of Income Dynamics)는 미국의 일반가구(가족) 및 가구원을 대상으로 부와 빈곤의 동적 변화를 파악하기 위해 1968년부터 시작된 가구 단위의 패널조사로 가장 오래된 패널조사이다. PSID의 표본은 1966/1967년에 약 3만 가구를 대상으로 조사한 경제조사(Survey of Economic Opportunity : SEO)에 근거하여 약 2,000가구의 저소득층 SEO 가구를 부표본으로 추출하고, 또한 미국 전체 가구를 포함하

는 표본틀로부터 횡단 모집단을 대표하도록 표본을 추가하여 저소득층 뿐만 아니라 저소득층이 아닌 가구를 표본으로 추출함으로써 미국의 전체 가구에 대한 대표성을 확보하는 원표본을 구성한다. 그 후, 매년 가구 분리로 인해 원가구원이 이동하더라도 원표본가구의 모든 가구원을 추적하고, 출생 이후 적격 대상이 되면 표본에 포함하게 됨으로써 PSID의 표본은 시간이 흘러가면서도 미국 가구 및 가구원에 대한 대표성을 유지하도록 하고 있다.

1968년 미국 가족에 대한 부와 빈곤의 동적 변화를 파악하기 위해 추출한 국가적인 확률표본인 PSID의 원표본은 독립적인 두 프레임을 사용하여 추출된 표본을 결합하여 구성한다. PSID의 원표본은 SEO 표본가구로부터 저소득층을 과대표집하여 구성한 1,872가구(SEO표본)와 미국 일반가구의 추출틀로부터 추출한 2,930가구(SRC표본)를 결합하여 구성한다. 1차 조사에서의 가구원(개인) 표본 규모는 SRC 가구원이 9,481명, SEO 가구원이 8,749명으로 모두 18,230명이다. 1차 조사에서 저소득층을 과대표집한 이유는 1960대 후반 주된 저소득계층인 아프리카계 미국인 가족의 표본 규모를 고려하였기 때문이다.

PSID의 개인을 추적하는 기준은 횡단 시점뿐만 아니라 시간 변화에 따라서도 미국 전체 가구(가족)의 대표성을 유지하는 표본을 구성하기 위해 제안된 것이다. PSID의 개인(가구원) 원표본은 1차 조사 시점인 1968년도 조사에 PSID 가구에 거주하는 모든 가구원과 표본 가구원에 의해 출생 혹은 입양된 가구원도 포함하며, 개인 원표본이 가구에서 거주하지 않더라도 추적조사를 한다. 또한 가구에 대해 추적한 결과로부터 매년 변화된 가구 표본 및 인구의 가족 구성 활동(family-building activity)을 통한 분가가구를 구성하여 표본으로 포함한다. PSID에는 비원가구원표본도 포함하는데, 1968년 이후 결혼한 원가구원의 배우자로 구성하며, 원가구원과 함께 생활하는 동안에만 추적한다. 1992년 조사에서는 1991년 이전 무응답 원가구원은 재접촉하여 접촉된 원가구원으로 포함하였고, 1992년 이전에는 18세 이하 가구원은 원가구 혹은 원가구원과 함께 생활하지 않는 가구원은 추적하지 않았으나 1993년 이후에는 청년 가구원으로 추적하도록 한다. 이때, 18세 이하의 원가구원과 함께 생활하는 비원가구원

중 성인은 동거하는 기간에는 조사하여 PSID의 표본으로 추가해 별도 관리한다.

기본적으로 PSID의 최초 조사설계에서는 1968년 이후 변화된 가구 및 가구원 모집단에 대해서도 대표성을 유지하도록 표본을 확보하기 위해 매 조사시점마다 출생 코호트 표본을 추가하도록 하였지만 1968년 이후 변화된 가구 및 가구원을 적절히 대표하지 못하게 된다. 이는 1968년 이후 이주한 이민자 중 PSID의 원가구와 함께 거주하는 이민자는 추적되지만 표본 가구원과 함께 거주하지 않는 이민자는 추적을 하지 않기 때문이다.

첫 번째는 1968년 이후 이주한 이민자 그룹을 표본으로 추가하기 위해 1990년에 멕시코, 푸에르토리코, 쿠바 등 라틴계 가구 2,043가구를 표본으로 추가해 확대한다. 이 표본 확대는 가장 많이 이민을 오는 3개국의 이민자에 대한 대표성은 확보하였지만 1968년 이후 이주한 모든 이민자에 대한 대표성은 확보하지 못하는 한계를 지닌다. 이러한 단점으로 인해 이 표본은 1995년도에 중지한다. 참고로 1995년도 중지하는 시점에서 1,834가구를 유지하고 있었다.

두 번째는 1997/1999년에 1968년 이후 1997년까지의 이민자그룹을 대표하는 표본을 추가하기 위해 약 500가구(가족)를 추출하였으나 조사에 참여한 441가구만 표본에 추가하여 PSID 표본을 갱신한다. 이 표본에는 1968년 이후 미국으로 이민 온 이민자, 1969년 이후 출생한 아동, 1968년 1차 조사 시점에 미국에 거주하지 않았던 가구를 포함하고 있으며, 이 표본을 1997 PSID 이민자 표본(1997 PSID Immigrant Refresher Samples)이라고 부른다. 즉 1997년도에 1968년 이후 미국으로 이주한 이민자가 증가하고, 비원가구원이 원가구로 유입되는 상황이 많아지면서 미국 가구의 구성 및 1968년 이후의 변화된 상황을 반영하고, 최근 모집단의 대표성을 확보하도록 유입 및 이민자 표본을 추가한 것이다. 1997 이민자표본의 원가구원에 대해서도 PSID의 추적 원칙을 동일하게 적용하여 분가된 이민자 가구의 원가구원은 계속 추적하여 표본에 추가한다.

세 번째는 자녀의 결혼, 이혼 등 원가구 및 원가구원의 분가로 인해 원가구가 약 5,000가구로 증가하면서 조사비용이 증가하고 추적조사하는 데

어려움이 발생하여 SRC와 SEO의 원가구를 1996년 8,511가구에서 1997년 6,303가구로 축소한다. 실제로 SEO 표본은 분가로 인한 가구의 자연적인 증가가 많았기 때문에 표본가구는 주로 SEO 표본을 감소시키는데 3,964가구에서 1,714가구로 2,250가구를 감소시킨다. SEO 표본은 주로 저소득층인 아프리카계 가구이므로 경제적 전이, 빈곤 및 건강 등 다른 요인에 대한 불공정한 차이에 관한 연구를 위해 계속 유지하고 있다.

PSID의 현재(2013년 기준) 표본 규모를 보면, 가구는 SRC 표본이 5,450가구, SEO 표본이 2,932가구, 이민자 표본이 681가구로 모두 9,063가구이고, 가구원(개인)은 SRC 표본이 14,562명, SEO 표본이 8,099명, 이민자 표본이 2,291명으로 모두 24,952명이다.

2. 영국 BHPS와 UKHLS

가. BHPS

영국의 BHPS(British Household Panel Survey)는 1991년부터 시작된 조사로, 현재는 일반가구를 대표하도록 추출된 표본가구의 성인(16세 이상) 가구원을 대상으로 매년 추적조사하는 패널조사이다. BHPS의 표본 추출방법은 2단계 층화집락계통추출법이다. BHPS의 주된 목적은 영국 내 가구 및 개인(가구원) 수준에서 사회-경제적 변수들의 변화, 원인 및 결과에 대해 식별하고, 모형화하고, 예측하는 등의 과정을 통해 장래의 사회 경제적 변화를 파악하는 것이다. 조사는 기본적으로 사회과학분야의 연구를 위한 자료를 제공하면서도 다양한 영역과의 학제 간 연구를 지원하기 위해 설계된 것이다.

표준의 BHPS 표본 관리 규칙은 1998년 제정하였으며, 현재의 표본 관리 및 추적 조사 원칙은 다음과 같다.

첫째, 1차 조사에서 가구원(개인)의 면접여부에 관계없이 조사된 가구에 거주하는 1명 이상의 성인이 있는 경우는 개인 표본으로 관리하여 추적조사를 수행하는데, 2차 조사 이후 조사된 시점에서 새로운 개인(가구원)의 유입으로 간주한다.

둘째, 독립, 이혼 등의 사유로 원가구로부터 분가한 가구가 새로운 가구를 형성하는 경우는 새로운 가구의 모든 성인을 신규 유입된 개인 표본으로 추가하는데, 기본적으로 1년 중 6개월 이상 연속하여 함께 거주해야 한다.

셋째, 원가구원으로부터 출생하거나 혹은 신규 유입된 가구원 중 16세 미만의 아동은 16세부터 표본으로 추가한다. 참고로, 4~7차에 11~15세 아동을 대상으로 조사(BYP)를 수행하기도 했다.

그 결과, 각 시점에는 원표본가구원, 원가구원의 식구 진입 가구원, 그리고 새로 유입된 성인으로 구성되어 현재 BHPS의 표본은 1990년대 이후 변화한 영국의 인구(개인) 모집단을 대표하도록 유지하고 있다.

BHPS의 표본 규모는 1차 조사에서는 영국(Britain)의 250개 지역으로부터 추출된 약 5,500가구의 가구원(개인) 10,300명이다. 1차 조사 이후 모집단 대표성 확보를 위한 BHPS의 표본 추가 및 확대에 대한 현황을 보면 다음과 같다.

〈표 2-2〉 BHPS의 표본 추가 현황과 규모

	추가 사유	특징
1991	잉글랜드 지역	최초 패널
1997	저소득층 패널 추가	UK ECHP의 부표본 (2001년 ECHP 중지)
1998	표준 BHPS 규칙 확정	
1999	스코틀랜드(칼레도니아의 북부와 서부)/웨일스 표본 추가	과대표집/지역 간 비교 (목표 표본 : 각각 1,500가구)
2001	노던아일랜드 확대 패널(NIHPS)	별도 패널관리/전국 대표성 확보 (목표 표본 : 2,000가구)
2009	UKHLS 1차 조사 실시	UK 가구 종단 조사 실시
2010	UKHLS 2차 조사에 BHPS를 포함	BHPS 원표본 중 동의 표본만 포함

첫째, 1997년(7차 조사)에 ECHP(European Community Household Panel) 부표본을 위한 저소득층 표본을 추가하여 운영하다가 2001년(11차 조사)에 중지하였고, 둘째, 1999년(9차 조사)에 지역별 개별 분석 및 지역 간

비교를 위해 상대적으로 표본 규모가 작은 스코틀랜드와 웨일스 지역의 표본에서 각각 1,500가구를 추가하였고, 셋째, 2001년(11차 조사)에는 영국 전역(UK-wide)으로 확대하기 위해 노던아일랜드 표본을 2,000가구 추가하여 영국 전체의 대표성을 확보하였다. 노던아일랜드 표본은 BHPS에 포함되어 있지만 별도의 패널조사로 관리하고 있다. 마지막으로, 현재 BHPS는 2009년 새로이 계획된 영국 가구 종단조사(UKHLS)의 2차 조사(2010~2011년)에 포함되어 조사가 수행되고 있다. 이때 BHPS 18차 조사에 참여한 표본 중 잉글랜드, 스코틀랜드, 웨일스 및 노던아일랜드 지역에서 원가구원으로 구축한 개인 원표본과 그 자녀 중 UKHLS에 동의한 표본만을 UKHLS에 포함하였다.

나. UKHLS

UKHLS(UK Household Longitudinal Survey)는 사회 및 경제 연구에 필요한 자료를 제공하는 패널조사 중 전 세계에서 가장 규모가 큰 패널조사이다.

UKHLS는 2009년 1월 처음 시작해 원가구의 가구원(개인)을 매년 추적조사하여 자료를 수집하는데, 조사기간은 24개월이고, 각 개인으로부터 매년 자료를 수집한다. UKHLS는 영국의 모든 유형의 가구 및 가구원을 조사함으로써 모든 연령대를 포함하여 전망적(prospective) 관점이 가능한 자료를 수집하고, 장기적인 직업의 선택과 이동에 영향을 주는 건강, 웰빙, 가정 환경 및 개인적인 관계 등 사회, 건강 및 생명 과학을 연계하는 다양하고 많은 영역의 연구를 위한 자료를 수집해 제공한다. 특히, UKHLS에서는 다양한 사회에서 유전자 구성과 관련하여 건강에 영향을 주는 환경을 살펴보기 위해 2만 명의 생의학자료를 수집해 제공하기도 하였다(2차 및 3차 조사).

UKHLS의 표본은 기본적으로 영국의 일반 가구 및 가구원(개인)을 대표하도록 추출되어 관리하고 있다. 표본 규모는 약 4만 가구, 10만 명의 개인으로 구성되어 있는데, 원표본(Original Samples Members : OSM)은 다섯 가지 유형으로 관리되고 있다.

첫 번째 유형은 설계 시점(2009~2010년, 1차 조사)에서 영국(잉글랜드, 스코틀랜드, 웨일스 및 노던아일랜드)의 일반가구 및 가구원 모집단을 대표하도록 추출한 신규 대규모 표본(General Population Samples : GPS)으로, 잉글랜드, 스코틀랜드 및 웨일스 지역은 47,520표본, 노던아일랜드는 2,395표본으로 모두 49,915표본이며, 이를 UKHLKS의 1차 조사에서 구성한 원표본구성원으로 정의한다.

두 번째 유형은 일반 모집단(GPS)과의 비교를 위한 표본(General Population Comparison Samples : GPC)으로, GPS의 2,640개 psu(sector로 사용) 중 40%에 해당하는 psu에서 하나의 표본을 확률 추출한 GPS의 부표본으로 구성되어 있다. 이 GPC 표본은 GPS 표본에 포함되어 분석되는 표본이다. 비교 표본에는 출생코호트 비교연구를 위한 16세 미만의 개인이 1.3만 명, 바이오뱅크와의 비교연구를 위한 40~69세 개인이 3.5만 명, ELSA와의 비교연구를 위한 1.5만 명의 표본이 포함되어 있다.

세 번째 유형은 민족성을 포함해 소수민족에 대한 부스터표본(일종의 과대표집)으로, 최소 1,000명 이상의 성인을 포함한 영국 최초의 종단조사 중 하나이다. 이를 위해 인도계, 파키스탄계, 방글라데시계, 캐리비안계 및 아프리카계(다른 소수민족 일부 포함) 등 5개의 소수민족을 고려해 표본(Ethnic Minority Booster samples : EMB)을 별도로 관리하지만 영국을 대표하는 원표본구성원(OSM)의 표본으로 본다. EMB의 표본추출 과정은 OSM과 기본적으로 동일하다. 표본은 2001 센서스와 매년 수행되는 인구조사(Annual Population Survey)에 근거하여 표본 규모를 조정한다. EMB의 표본 추출을 위해 소수민족이 거주하는 3,145개 psu(5개 소수민족 인구의 82~93%가 포함된 것으로 추정)를 소수민족의 인구수에 따라 4개 층으로 층화하여 최소 3가구 이상 포함된 층에서는 추출률을 1/4, 1/8, 1/16 등으로 하여 표본 psu를 추출한다. 이때 표본가구는 psu 내에서 적격 가구가 최소 2가구 이하가 되도록 제한하여 771개 psu를 추출하였는데, 스코틀랜드 6개, 웨일스 7개, 잉글랜드 758개이다. 잉글랜드 중 런던은 412개로 제한하였다. 5개 소수민족별 표본은 통계적 효율을 고려하여 소수민족별 목표 표본규모를 결정하였으며, GPS의 psu와 마찬가지로 계통추출로 psu를 추출하였다. 표본가구 및 가구원 추출 과정은 ① 표본

psu마다 추출률을 이용하여 추출된 표본가구마다 목표 민족그룹의 가구원(개인)이 거주하는지 먼저 확인하고, ② 민족 그룹별 목표 표본크기의 표본가구원을 확보하기 위해 전수추출(인도혼혈계, 방글라데시계, 캐리비안혼혈계, 스리랑카계, 중국계, 터키계)하거나 적절한 확률적 메커니즘을 이용하여 목표 비율만큼 표본추출(인도계, 파키스탄계, 캐리비안계, 아프리카계, 극동아시아 및 중앙아시아계)하여 소수민족을 위한 표본(EMB)을 구축하였다. ③ 추출된 표본가구에서 목표 소수민족의 모든 가구원은 EMB 표본으로 선정하지만 다른 소수민족의 개인은 EMS 표본으로 관리하지 않으면서 EMB 가구원과 함께 거주하는 동안에 한하여 임시 표본(Temporary Sample Members : TSM)으로 관리하여 자료를 수집하고 있다.

네 번째 유형은 UKHLS 2차 조사(2010~2011년)에 포함된 BHPS의 18차 조사에 참여한 적격 원표본 중 UKHLS에 동의한 표본이다. BHPS 중 UKHLS에 포함된 표본은 잉글랜드 지역의 표본(1991년), 스코틀랜드와 웨일스의 부스터 표본(1999년), 노던아일랜드 지역의 표본(2001년)만 포함하고 있다. BHPS는 UKHLS에 포함된 시점에서는 각 지역마다 BHPS의 표본에 추가된 시점 이후 원가구원인 부모로부터 출생된 자녀도 포함하지만 2차 이후는 모로부터 출생된 자녀만 포함하고 있다. UKHLS에 포함된 BHPS의 규모는 2차 조사(2010~2011년)에서 약 8,000가구 중 6,700가구이다.

다섯 번째 유형은 UKHLS의 본조사에 앞서 해마다 새로운 문항의 질 평가, 응답률 개선 방법, 모드 효과와 같이 조사방법론의 주요한 이슈에 대한 연구를 위해 관리하는 혁신패널(Innovation Panel)을 운영하고 있다. 혁신 패널은 1,500가구이다.

결론적으로 UKHLS의 표본은 원표본구성원 표본(OSM), 임시 표본(TSM), 영구 표본(Permanent Samples Member : PSM) 등 세 가지 유형으로 운영된다. 원표본구성원 표본(OSM)에는 GPS, GPC, EMB, BHPS 등 원표본 가구원과 그 자녀들이 포함되고, 임시 표본(TSM)에는 EMB 표본에 속하지 않는 개인, OSM 및 PSM 가구에 함께 거주하는 OSM 표본이 아닌 신규 유입 가구원 등이 포함되고, 영구 표본(PSM)은 OSM과 함께 생활하지 않는 임시 표본이다. OSM은 매년 추적하지만 TSM과

PSM은 OSM 가구에 함께 거주하거나 적격한 대상으로 포함될 때만 추적한다.

3. 독일의 SOEP

독일의 SOEP(Social-Economic Panel Survey)는 독일의 전체 거주자를 대표하는 패널조사로, 모집단은 독일 전 지역의 일반가구(the private household)에 거주하는 거주자 모집단(the residential population)이며, 조사대상은 표본가구를 추출단위로 하여 표본가구 내에서는 16세 이상 가구원(거주자) 전체가 대상이다. SOEP의 주된 목적은 가구 구성, 직업, 고용, 소득, 건강 및 생활 만족도 등 많은 분야의 연구를 위한 자료를 수집해 제공하는 데 있으며, 전 세계 패널조사 자료와도 공유하고 있다.

가. SOEP의 표본 확대

SOEP는 독일 통일 전인 1984년에 서독 지역에 거주하는 거주자 표본(표본 A)과 당시 외국에서 독일로 이주한 외국인을 과대표집한 외국인 표본(표본 B)을 추출하여 처음 조사가 시작되었다. 외국인 표본 추출을 위한 국가는 터키, 이탈리아, 스페인, 그리스 및 유고이다. 1990년 통일 후 독일 경계가 변경(1990년)되고, 이민, 입양 등으로 거주자 모집단이 변화되어 표본의 대표성을 유지하기 위해 표본을 추가해 확대 개편하게 되는데, 표본을 추가할 때 최초 설계틀을 유지하도록 표본추출구조를 동일하게 적용한다. 일부 표본 그룹에 대해서는 통계적 효율을 높이기 위해 과대표집으로 추가해 확대 표본을 구축하고 있다. 그 후 SOEP는 오랜 기간 동안 변화하는 거주자모집단의 특성을 반영하고, 표본의 대표성을 확보하기 위해 지속적으로 표본을 확대 개편하고 있는데, 효과적인 표본 관리를 위해 SOEP 표본 구성 및 개편 시점을 구분하여 A부터 일련번호를 부여해 관리한다.

SOEP 표본의 개편 내용과 규모를 요약해 정리하면 다음과 같다.

1) 우선, 1990년 6월 통일이 되어 독일의 경계가 동독 지역을 포함하게

되므로 SOEP는 독일 전 지역 거주자 모집단의 대표성을 확보하기 위해 동독 지역의 거주자를 추출하여 표본을 확대한다(표본 C). 추가된 표본 규모는 2,179가구, 4,453명의 거주자이다.

2) 1994/1995년에는 1984년 이후 독일로 이주한 이민자가 증가하는데, 특히 소비에트 국가로부터 이주한 게르만계를 고려하기 위해 이민자 표본을 확대한다. 이민자 표본은 1994년에 1차로 표본(표본 D1 : 236가구, 471명의 거주자)을 확대하고, 1995년에 2차로 표본(표본 D2 : 295가구, 607명 거주자)을 확대해 1994/1995년에 이민자 표본을 추가해 확대 개편한다. 1994/1995년에 확대된 이민자 표본의 확대 규모(D1/D2)는 모두 531가구, 1,078명의 거주자이다.

〈표 2-3〉 SOEP 표본유형별 추가표본 규모

		연도	가구	개인	거주자 (응답자)	아동
A	서독 거주자	1984	4,528	11,422	9,076	2,290
B	외국인	1984	1,393	4,830	3,169	1,638
C	동독 거주자	1990	2,179	6,131	4,453	1,591
D1	이민자	1994	236	733	471	248
D1/D2	이민자	1995	531	1,668	1,078	517
E	확대표본 1	1998	1,067	2,446	1,923	466
F	확대표본 2(가구추출)	2000	6,043	14,510	10,880	2,991
G	고소득자	2002	1,224	3,538	2,671	693
H	확대표본 3	2006	1,506	3,407	2,616	623
I	혁신(인센티브)	2009	1,531 (1,495)	3,428	2,509 (2,432)	620
J	확대표본 4	2011	3,136	6,873	5,161	1,147
K	확대표본 5	2012	1,526	3,286	2,473	563
L1	가족패널 FiD 1	2010	2,074	7,939	3,770	3,900
L2	가족패널 FiD 2	2010	2,500	9,063	4,227	4,611
L3	가족패널 FiD 3	2011	924	3,645	1,487	2,092
M1	이민자 고용표본 1 (IAB-SOEP)	2013	2,723	8,522	4,964	2,481
M2	이민자 고용표본 2 (IAB-SOEP)	2015	1,096	3,048	1,711	927

주 : http://about.paneldata.org/soep/dtc/sample.html

3) 장기간의 조사로 인한 표본 마모와 변화된 모집단 특성을 반영해 전체 거주자모집단의 대표성을 높이기 위해 표본크기를 증가시키는 개편을 1998년(표본 E : 1,067가구, 1,923명 거주자)과 2000년(표본 F)에 수행한다. 특히, 2000년에는 표본가구를 기존 표본과 독립적으로 추출하여 16세 이상 모든 거주자를 표본으로 추가하는 방식으로 확대 개편한다. 2000년에 추가된 표본 규모는 6,043가구, 10,880명의 거주자이다.

4) 2002년에 월 가구소득이 7,500DM 이상인 가구를 독립적으로 추출하여 사전 스크리닝 과정으로 고소득층 표본을 과대표집하여 표본(표본 G)을 확대한다. 2003년에는 가구소득을 4,500유로 이상인 가구로 기준을 변경한다. 2002년에 추가된 고소득층의 표본 규모는 1,224가구, 2,671명의 거주자이다.

5) 2006년에 변화된 모집단 특성을 반영하여 모든 거주자 가구(all residential household)를 포함하도록 표본을 확대한다(표본 H). 2006년 확대된 표본 규모는 1,506가구, 2,616명의 거주자이다.

6) 2009년에도 모집단 대표성을 유지하기 위해 기본 SOEP표본과는 독립적으로 1,531가구(2,509명 거주자)를 추가하여 표본을 확대한다(표본 I). 2009년에 자료수집방법의 개선, 새로운 조사 문항 개발 및 패널의 참여율 제고 방안을 검증하기 위해 혁신 표본(인센티브 표본)을 추출하였으며, 2012년에 2012년의 자료 분포를 고려해 표본 I로부터 추출한 부표본으로 SOEP 혁신 표본(SOEP-IS)을 구축한다. 2012년에 구축한 SOEP 혁신 표본(SOEP-IS)의 규모는 1,040가구(2,113명 거주자)이다.

7) 2011년과 2012년에 변화된 독일 거주자모집단의 대표성을 확보하고, 또한 SOEP의 전체 표본규모를 확대하기 위해 표본을 추가한다. 2011년에 확대된 표본 규모(표본 J)는 3,136가구, 5,161명의 거주자이고, 2012년에 확대된 표본 규모(표본 K)는 1,526가구, 2,473명의 거주자이다.

8) 2013년과 2015년에는 이민자 모집단의 대표성을 보완하기 위해 표본을 확대한다. 2013년에는 연방 고용부에 등록된 정보를 이용하여 2,723명을 추출하여 추출된 표본의 가구와 가구에 함께 거주하는 4,964명의 거주자를 표본으로 추가한다(표본 M1). 2015년에도 이민자가 증가하여 모집단이 변화함에 따라 2013년 추출 과정과 동일한 방법으로 1,096가구,

1,711명의 거주자를 추가한다(표본 M2). 2013년과 2015년의 확대 표본은 IAB-SEOP표본으로 관리한다.

9) 2014년에는 독일의 가족을 포괄하고, 가족 연구와 관련한 자료를 제공하기 위해 2010년부터 SOEP와 별도로 수행되는 가족 패널을 SOEP 표본으로 병합하여 표본을 확대한다(표본 L). 가족표본(FiD)은 가족을 표본으로 추출하여 SOEP와는 별도로 진행되는 조사로 2010년 처음 시작한 조사이다. 2010년 설계 시점에서 2,074가구, 3,770명의 가구원을 표본(표본 L1)으로 구성하였으나 조사 완료 후 2010년에 다시 2,500가구 4,227명의 가구원을 추가(표본 L2)하게 되고, 2011년에 다시 924가구, 1,487명의 가구원을 추가(표본 L3)해 가족표본을 확대한다.

나. SOEP-IS의 표본 확대

SOEP의 혁신 패널(SOEP-IS)은 2009년에 SOEP 패널에 대한 자료수집방법의 개선, 새로운 조사 문항 개발 및 패널의 참여율 제고 방안을 검증하기 위해 혁신(혹은 인센티브) 표본을 구축한다. SOEP-IS는 연구자들에게 단기간의 실험, 특별한 주제에 대한 설문이나 SOEP의 핵심 연구와는 별도의 주제에 대한 장기간의 설문을 검정하는 데 유용하게 사용되는 표본이다.

SOEP-IS의 구축 및 확대 과정을 요약하면 다음과 같다.

1) 2008년 이전에는 1998년에 구축한 표본 E로부터 373가구(963명 거주자)를 추출하여 혁신 표본을 구축하고,

2) 2009년에는 SOEP의 표본 I로부터 1,495가구(3,052명 거주자)를 추출하여 혁신 표본을 확대하는데, 2009년에 추출된 혁신 패널은 2010년과 2011년의 처음 2년간은 주요한 자료 분포 특성을 유지하고 있었다. 2011년에 조사한 1,040가구(2,113명 거주자)를 SOEP 혁신 표본(SOEP-IS)으로 명칭을 부여해 관리한다.

3) 2012년, 2013년, 2014년에도 계속하여 추가하여 SOEP 혁신 표본(SOEP-IS)도 확대하고 있다.

그 결과, 2014년의 SOEP 혁신 표본(SOEP-IS) 규모는 3,716가구, 7,322

명의 거주자(개인)로 운영된다.

〈표 2-4〉 SOEP 혁신 표본(SOEP-IS) 표본규모

	1998 ~2000	2009	2010	2011	2012	2013	2014
표본 E	373(963) SOEP	447(934) SOEP	453(936) SOEP	464(944) SOEP	339(642) SOEP-IS	310(599) SOEP-IS	295(580) SOEP-IS
표본 I		1,495 (3,052) SOEP	1,175 (2,450) SOEP	1,040 (2,113) SOEP-IS	928 (1,826) SOEP-IS	864 (1,728) SOEP-IS	797 (1,596) SOEP-IS
2012표본					1,010 (2,005) SOEP-IS	833 (1,688) SOEP-IS	772 (1,611) SOEP-IS
2013표본						1,166 (2,141) SOEP-IS	928 (1,870) SOEP-IS
2014표본							924 (1,665) SOEP-IS
계 : 가구 (개인)	373 (963)	1,622 (3,986)	1,628 (3,386)	1,504 (3,057)	2,277 (4,473)	3,173 (6,163)	3,716 (7,322)

4. 호주의 HILDA

가. 조사 개요 및 특징

호주의 HILDA(Household, Income and Labor Dynamics in Australia) 조사는 2001년부터 시행된 가구 단위의 패널 조사이다. 주요한 목적은 소득 변화, 노동시장 변화 및 가족 변화 등 세 가지 영역과 관련된 분야의 연구에 필요한 기초 자료를 제공하기 위함이다. 측정 자료는 가구 및 가족 관계, 아동 보호, 고용. 교육, 소득, 소비, 건강과 웰빙, 다양한 주제뿐만 아니라 삶 전반에 대한 사건과 경험 등에 대한 태도와 가치 등을 포함한 호주 내 일반 사람들의 삶 전반에 대한 자료 등이다. HILDA는 빈곤의 발생과 지속, 은퇴 전환기의 자산과 수입, 신체와 정신 건강의 관계 및 생

애주기별 변화의 영향, 부와 행복의 국제 간 비교 연구 등 경제, 사회과학 및 사회정책 분야의 다양한 주제들의 국내외 연구에 폭넓게 활용된다.

대상 모집단은 호주의 일반주택에 거주하는 개인(가구원)으로 정의하고, 호주에서 근무하는 외국 외교인력, 해외에 장기 거주하는 개인, 조사 및 접촉이 어려운 지역에 살고 있는 사람들을 제외한 호주 내 일반주택에서 생활하는 가구의 15세 이상 모든 가구원(개인)이다.

HILDA 표본은 지역기반의 3단계 층화집락추출법으로 표본을 추출하였다. 주 및 대도시 등의 층에서 1996년 총조사의 총조사구(Census Collection Districts : CCD)를 psu로 하여 주거여부에 관계없이 주택을 계통추출하고, 주택 내에서 최대 3가구 이하의 표본가구를 추출하여 조사 참여여부에 대한 스크리닝 과정을 통해 1차 조사의 표본을 구축한다. 1차 조사에서 패널 구축을 위해 추출된 적격표본가구는 모두 11,693가구이며, 조사 참여에 동의한 7,682가구를 원표본가구로 구축한다.

HILDA는 호주 내 일반주택 내 가구에서 생활하는 인구를 대표하도록 가중치를 부여한다.

확대표본의 가중치는 1차 조사부터 시행한 원표본은 원표본 가중치를 부여하고, 2011년에 추가된 추가표본은 원표본의 가중치 산출 과정을 동일하게 적용하여 추가표본의 가중치를 계산하여 원표본과 추가표본을 결합한 확대표본에 대한 가중치를 산출한다.

HILDA의 주요한 특징을 요약하면, ① 경제, 주관적 웰빙, 노동시장의 변화, 가족 변화 등에 대한 자료를 수집하고, ② 각 시점마다 부, 은퇴 등 특정 주제에 대한 설문 모듈을 채택하고, ③ 1차 조사에서 7,682가구의 19,914명의 개인을 대상으로 시작하여 조사를 수행하다가 2011년(11차)에 2,153가구, 5,477명의 표본을 추가해 패널을 확대하고, ④ 표본가구의 15세 이상 모든 구성원에 대해 자료를 수집하는데 표본가구의 15세 미만 아동은 15세 이후 표본으로 포함하며, 표본으로 포함된 개인은 계속 추적조사한다.

HILDA 자료는 한국, 호주, 캐나다, 독일, 영국, 스위스, 러시아 및 미국의 인구 패널 자료를 통합한 국가별 횡단 파일(the Cross-National Equivalent File : CNEF)에 제공된다.

나. 2011년의 추가표본 개요

2001년 이후 이민자 증가, 출생 및 사망, 가구의 분가 및 통합 등 가구 및 개인 단위의 모집단이 변화하고, 동시에 1차 조사 이후 원표본에서도 조사거부로 인한 이탈, 추적실패, 해외이주 등 표본 마모가 발생하여 원표본의 구조도 모집단 구조와 차이를 보이므로 2011년에 표본을 추가하여 확대한다. 특히, 마모되는 표본은 유지 표본과 체계적으로 다른 특성을 갖기 때문에 비확률 변동에 심각하게 영향을 주게 된다. 이러한 한계를 보완하기 위해 HILDA에서 2011년에 추가표본을 1차 조사와 유사한 과정으로 표본 설계를 하여 3,117적격가구를 추출하고, 조사참여에 동의한 2,153가구를 확대표본으로 추가한다. 2011년 추가표본 설계와 2001년 원표본 설계에 대한 차이를 요약하면, ① 2006년 총조사 지역을 기준하고, ② psu 추출을 위한 주택수는 실제 거주하는 주택수를 기준으로 하며, ③ 추가표본 규모가 작으므로 층화추출하지 않고 주, 지역 등의 순으로 정렬하여 계통추출한 점이 다르다. 두 설계에서 설계 가중치가 약간 다르지만 거의 차이가 없는 것으로 분석되었다.

일반적으로 가구 기준의 패널조사에서 가구 및 분가가구를 추적조사함으로써 가구원 분가와 독립, 해외 이주, 사망과 출생 등 가구 구조의 변화

〈표 2-5〉 HILDA 11차(2011년)의 표본특성별 표본 규모(가구와 개인)

		원표본(PSM)		추가표본
		2001	2011	2011
가구	대상 가구수	12,252	10,526	3,250
	적격 가구수	11,693	10,440	3,117
	응답 가구수	7,682	7,390	2,153
개인	총 규모	19,914	29,489	5,451
	PSM 규모	19,914	19,914	5,451
	+PSM 규모		3,482	
	TSM		2,529	
	TSM(제외)		3,564	
	적격 개인	15,127	14,352	4,280
	응답자	13,969	13,603	4,009
	무응답자	1,158	749	271
	아동	4,787	3,601	

를 파악하여 모집단의 변화에 대한 유용한 정보를 얻게 된다. HILDA 조사는 기본적으로 원표본(개인, Permanent Sample Member : PSM)은 해외 이주, 사망 등의 조사가 불가능하지 않으면 계속 추적할 뿐만 아니라 ① 원표본에 의한 출생과 입양된 개인, ② 원표본 개인의 출생과 입양아동의 비원가구원인 부모, ③ 2001년 이후 호주에 이주한 개인은 원표본의 유입표본(Convert PSM)으로 별도로 관리하여 계속 추적한다. 이들 개인을 제외한 가구에 거주하는 개인은 원표본과 거주하는 동안만 조사하는 임시 표본(Temporary Sample Member : TSM)으로 관리한다.

지금까지 KLIPS와 유사하게 일반가구 및 가구원을 대상으로 경제활동, 고용, 취업 등을 파악하기 위한 국가별 패널조사에서 표본을 추가하는 사례를 살펴보았다. 국가별 패널조사의 개편 양태를 요약해 보면, 독일이 가장 많은 개편이 있었고, 미국, 영국과 독일은 주로 조사지역 범위를 확대하거나 조사모집단으로 설정한 일반 국민의 범주에 외국인 혹은 이민자를 포함함으로써 인구모집단의 커버리지 확대를 위해 주기적으로 추가표본을 추출하여 확대 개편하고, 독일과 호주는 주된 사유가 변화된 모집단의 대표성 확보와 표본 마모로 인한 표본의 대표성 유지를 위해 현재 모집단에 표본을 추가하는 방법으로 표본을 확대 개편하고 있다. 이러한 외국의 추가표본 사례는 KLIPS의 표본 확대 방안을 결정하는 데 참고하는 것이 바람직하다.

제4절 한국노동패널조사(KLIPS)에서의 표본 확대 방안

1. 표본 확대 및 패널 관리 방안

가. KLIPS의 조사 개요와 표본 현황

1) 98표본의 설계 개요

한국노동패널조사(KLIPS)는 1998년에 당시 도시거주 가구의 노동시장

특성을 파악하기 위해 시작된 조사이다. 목표모집단은 1998년 현재 우리나라 도시거주 가구와 15세 이상 개인으로 정의하고, 대상모집단과 표본틀은 1995년 인구주택총조사의 10% 표본조사구인 전국의 21,938조사구 중 제주도와 군부지역을 제외한 도시지역 19,025조사구를 표본틀로 선정하여 표본을 추출하였다. 1차 조사에서의 표본추출방법은 1단계에서 1,000개의 조사구를 선정하고, 2단계에서 각 조사구 내에서는 고용구조특별조사의 조사대상 가구 중 5가구를 무작위 추출하는 2단계 층화집락계통추출법을 사용하여 원표본가구를 추출하고, 표본가구 내 15세 이상의 가구원과 동거인 전체를 표본개인으로 추출하였다. 그러나 당시 예산상의 제약문제로 1997년 통계청이 수행한 고용구조특별조사의 조사가구가 선정되도록 매칭하였다. 1차 조사에서 조사에 동의한 5,000가구를 표본가구로 선정하여 원표본가구로 정의한다. 표본가구 내 15세 이상 원표본가구원은 13,321명이다.

2) 2009 추가표본의 설계 개요

한국노동패널조사(KLIPS)는 원표본 내 원표본가구원의 독립, 원가구원의 일부 이주 및 분가 등에 관계없이 추적이 가능하면 계속 추적조사하는 것을 추적 원칙으로 설정하고, 또한 원가구원에 의한 출생 및 원가구원의 사망 등도 반영함으로써 모집단 규모 및 구조 변화를 고려해 대표성을 유지하도록 하고 있다. 하지만 98년 표본은 도시거주자를 대상으로 하였으나 일부 패널이 농촌지역으로 이주하고, 또한 농촌지역을 포함한 전체 모집단에 대한 분석 요구가 증가하고, 그리고 장기간 관리된 원표본에 대한 추적 실패 및 참여 거부 등 패널 이탈로 인한 표본 마모가 발생하여 표본의 구조와 모집단 구조가 괴리를 보임으로써 표본의 추가가 요구되었다.

12차(2009년) 조사시점에서 패널 이탈에 의한 표본 마모와 도시 지역 표집에 따른 대표성의 한계를 극복하고자 전국 단위의 가구로 모집단을 확장하게 된다. 추가표본을 위한 모집단은 2009년 3월 현재 대한민국 영토(제주도 포함) 내의 일반가구에 거주하고 있는 15세 이상 개인으로 정의하고, 표본추출틀로는 2005년 인구주택총조사의 조사구로 설정한다. 추

가표본 규모는 2005년 인구주택총조사 모집단과 KLIPS 표본의 동부와 읍면부의 분포를 동일하게 맞추기 위해서 동부에서 900가구, 읍면부에서 600가구 등 모두 1,500가구를 추가표본 규모로 결정해 1,415가구를 표본으로 추가한다.

추가표본의 표본추출방법은 동부와 읍면부에서 다르게 적용하는데, 동부에서는 기존 자료의 가구원수 분포와 주택유형별 분포가 센서스의 분포와 지나치게 차이가 나므로 이를 보정하는 방법을 별도로 사용한다. 먼저 모집단에서 조사구를 추출할 때 동부는 2005년 센서스 자료의 1~2인 가구수 및 단독주택 가구수 등 조사구별 특성치에 비례하도록, 읍면부는 조사구별 가구수에 비례하는 확률비례추출(Probability Proportionate Sampling) 방법으로 표본조사구를 추출한다. 추출된 표본조사구에서 계통추출법을 사용하여 각 조사구당 5개의 표본가구를 추출하여 스크리닝 과정을 통해 1,415가구를 추가표본으로 확정한 후 원표본과 결합하여 통합표본(the integrated sample)을 구축한다. 2009년에 추가표본을 추출하여 통합표본을 구축하는 시점에서 98표본은 98표본의 원표본 여부에 관계없이 2009년에 조사한 모든 가구 및 가구원을 통합표본의 원표본으로 추가한다. 2009년 조사에서는 98표본은 12차(2009년)조사로, 통합표본은 1차 조사로 구분하는데, 통합표본의 원표본은 6,721가구, 가구원 14,489명이다.

참고로, 통합표본은 98원표본과 별도로 구분하여 관리하지만 09통합표본은 기본적으로 98표본을 전국으로 확대한 표본이기에 98표본의 조사설계 틀을 동일하게 유지한다. 그래서 통합표본의 가중치는 98표본과 동일한 방식과 과정을 적용하는데, 2009년부터는 98표본 가중치와 통합표본 가중치를 별도로 구분해 제공한다.

3) 현행 표본 현황

2015년을 기준으로 보면, 98표본은 18차 시점이고, 09통합표본은 7차 시점이다. <표 2-6>은 2015년까지의 표본 현황을 요약한 표이다.

2015년 조사의 표본 현황을 표본별로 보면, ①98표본의 원표본가구 유지율은 2015년(18차) 조사에서 68.4%를 유지하고 있으며, ②09통합표본의 원표본가구(2009년에 응답한 모든 가구: 98표본의 원가구/분가가구 +

〈표 2-6〉 한국노동패널조사의 조사차수별 표본 현황

	전체 표본			98표본						통합표본					
	조사대상 가구	조사 성공	가구원 응답자수	조사대상 원가구-소멸가구	조사성공 가구수[1]	원표본 가구수[2]	원표본 유지율[3]	분가 가구	가구원 응답자수[4]	조사대상 원가구-소멸가구	조사 성공 가구수[5]	원표본 가구수[6]	원표본 유지율[7]	분가 가구	가구원 응답자수[8]
1차('98)	5,000			5,000	5,000	5,000	100.0	-	13,321						
2차('99)	5,000			5,000	4,507	4,378	87.6	129	12,037						
3차('00)	5,256			5,000	4,266	4,044	80.9	222	11,205						
4차('01)	5,435			5,000	4,248	3,866	77.3	382	11,051						
5차('02)	5,634			5,000	4,298	3,798	76.0	500	10,966						
6차('03)	5,839			5,000	4,592	3,862	77.2	730	11,541						
7차('04)	6,055			4,949	4,761	3,862	77.2	899	11,661						
8차('05)	6,262			4,935	4,849	3,822	76.4	1,027	11,580						
9차('06)	6,472			4,914	5,001	3,820	76.4	1,181	11,756						
10차('07)	6,730			4,899	5,069	3,775	75.5	1,294	11,855						
11차('08)	6,936			4,862	5,116	3,709	74.2	1,407	11,734						
12차('09)	8,602	6,721	14,489	4,832	5,306	3,657	73.1	1,649	11,739	6,721	6,721	6,721	100.0	-	14,489
13차('10)	8,763	6,683	14,118	4,802	5,379	3,606	72.1	1,773	11,582	6,694	6,398	6,232	92.7	166	13,641
14차('11)	9,059	6,686	13,900	4,786	5,404	3,528	70.6	1,876	11,377	6,676	6,388	6,082	90.5	306	13,410
15차('12)	9,228	6,753	14,000	4,770	5,469	3,517	70.3	1,952	11,444	6,641	6,434	6,016	89.5	418	13,427

〈표 2-6〉의 계속

	전체 표본			98표본						통합표본					
	조사대상 가구	조사 성공	가구원 응답자수	조사대상 원가구-소멸가구	조사성공 가구수[1]	원표본 가구수[2]	원표본 유지율[3]	분가 가구	가구원 응답자수[4]	조사대상 원가구-소멸가구	조사 성공 가구수[5]	원표본 가구수[6]	원표본 유지율[7]	분가 가구	가구원 응답자수[8]
16차('13)	9,389	6,785	13,888	4,741	5,501	3,472	69.4	2,029	11,331	6,597	6,457	5,904	87.8	553	13,303
17차('14)	9,600	6,838	13,169	4,741	5,552	3,451	69.0	2,101	10,757	6,589	6,493	5,840	86.9	653	12,595
18차('15)	9,692	6,934	14,013	4,701	5,632	3,421	68.4	2,211	11,446	6,530	6,577	5,793	86.2	784	13,373

주 : 1) '조사성공가구'는 98표본 원표본가구 중에서 실사가 성공적으로 이루어진 유효표본가구와 조사 성공한 분가가구를 포함한 총 조사성공가구임.

2) '원표본가구'는 1차년도 조사 당시 원표본이었던 5,000가구 중 해당 연도에 실사가 성공적으로 이루어진 가구임.

3) '원표본유지율'은 1차년도에 조사에 성공한 원표본 5,000가구 중에서 해당연도에 조사 성공한 원표본가구 비율임.

4) 15세 이상 가구원으로 면접에 성공한 가구원수.

5) '조사성공가구'는 통합표본 원표본가구 중에서 실사가 성공적으로 이루어진 유효표본가구와 조사 성공한 분가가구를 포함한 총 조사성공가구임.

6) '원표본가구'는 통합표본 1차년도(2009년) 조사 당시 원표본이었던 6,721가구 중 해당 연도에 실사가 성공적으로 이루어진 가구를 의미함.

7) '원표본유지율'은 통합표본 1차년도(2009년)에 조사에 성공한 원표본 6,721가구 중에서 해당 연도에 조사 성공한 원표본가구의 비율임.

8) 통합표본 중 15세 이상 가구원으로 면접에 성공한 가구원수.

09 추가표본의 원가구)의 유지율은 86.2%이고, ③ 2015년 조사에서 소멸가구를 제외한 조사대상 가구는 모두 9,692가구(98표본 8,157가구, 추가표본 1,535가구)이며, 조사에 성공한 가구는 6,934가구이고, 가구원은 모두 13,373명이다.

나. 표본 확대의 필요성

일반적으로 장기간 수행된 패널조사에서는 최초 표본설계의 대상 모집단에서 제외된 대상이나 지역 등에 대한 추가적인 자료 수집을 위해 모집단의 범위를 확장하거나, 설계 시점 이후 모집단 증가와 구조의 변화 혹은 표본 마모로 인해 표본 대표성 확보가 필요한 경우에는 현행 표본의 확대에 대한 검토가 필요하다.

현재 한국노동패널조사(KLIPS)는 1998년 도시지역을 대상으로 처음 시작된 이래 2009년(98표본 12차조사)에 농촌지역을 포함한 전 지역으로 범위를 확대하고, 장기간 추적조사를 수행함에 따른 패널 탈락, 참여 거부 증가 및 추적 실패 등의 표본 마모를 보완하고, 변화된 모집단 구조를 반영하기 위해 표본 규모를 증가시키는 추가표본을 추출하여 통합표본을 관리하고 있다. KLIPS는 패널을 대상으로 가구의 소득, 소비와 개인의 노동시장 활동을 장기간, 그리고 통합적으로 분석할 수 있는 자료를 제공하기 때문에 경제, 노동시장, 고용 등 사회과학의 여러 분야에서 연구 자료로 많이 활용되고 있으며, 특히 국가 간 연구의 기초자료로 제공되고 있다.

하지만 KLIPS를 이용하는 이용자들의 인구특성별 분석의 요구, 지역별 분석의 요구가 증가하고 있고, 또한 국내 조사 환경과 여건의 악화 및 패널 장기화에 따른 표본 마모의 지속적인 증가로 인해 표본구조가 변화하고 있으며, 1인 가구 및 다문화 가구 증가 등 변화된 모집단의 대표성 확보 및 유지가 요구되기 때문에 표본 확대를 검토해야 할 상황이다.

첫째, 현재 KLIPS의 표본은 약 6,500가구, 13,000여 명의 가구원으로 관리되지만 이를 성별, 연령대별 및 청년, 여성(특히 경력 단절 여성 등), 비정규직 등 세분화된 특정 계층의 분석을 위해서는 표본크기가 부족하

다. 즉 청년, 여성, 비정규직 및 1인 가구의 청년 등 특정 계층에 대한 모집단의 대표성을 확보하기 어렵고, 분석하더라도 작은 표본규모로 인해 표본오차가 증가하므로 작은 오차를 갖는 추정 및 분석이 어려운 실정이다. 예를 들면, 비정규직의 경우 현재 약 1,300여 명가량 사례수가 발견되는데 이들의 노동시장 이행을 성별, 연령별, 학력별로 구분하거나, 비정규직 유형(한시적 근로자, 시간제 근로자, 비전형 근로자)별로 분석을 할 경우 해당 사례수가 적어 추정오차가 커지게 되기에 분석결과를 신뢰하기 어렵게 된다. 따라서 노동시장에서 고용정책의 관심대상에 대한 분석이 가능하며 해당 계층의 대표성을 확보할 수 있는 표본 규모의 확대가 필요하다.

둘째, 노동 및 고용과 관련하여 가구 및 가구원 기준으로 분석하는 지역통계에 대한 수요가 증가하고 있으나 현재의 노동패널의 표본 규모로는 지역별 분석을 수행하기 어려운 실정이다. <표 2-7>에서 보듯이 98표본은 설계 시점에서 100가구 이상의 응답 표본가구로 시작하였으나 이사 및 표본 마모 등으로 인한 패널의 이탈이 지역별로 다르게 나타나, 일부 지역에서는 지역별 표본규모가 작아서 지역별 분석 및 통계를 생산하는데 한계가 있다. 현재 대전, 울산, 강원, 충남, 충북은 다른 지역보다 표본 이탈이 심해 해당 지역의 표본크기가 100개 미만으로 작아 지역통계 생산이 불가능하다.

셋째, <표 2-6>에서 보듯이 패널의 장기화에 따른 표본 이탈 및 마모의 문제, 분가가구의 추적 어려움 등으로 표본의 구조가 변화하여 표본의 대표성을 유지하는 데 어려움이 있으므로 표본의 대표성을 확보하기 위한 확대표본의 검토가 필요하다. 20대 청년이 다른 지역으로 분가한 경우는 추적의 어려움이 증가하고, 그로 인해 표본의 구조가 변화하게 되어 모집단 구조와 차이가 점점 커지게 되는 상황이므로 대표성 유지를 위해 표본의 확대가 필요하다.

넷째, 2015년 현재 모집단의 가구 및 인구 특성이, 09통합표본의 모집단인 2009년 모집단과 2016년 모집단을 비교해 보면, 1인 가구의 비율이 가장 많게 되며, 출생아수 감소 및 청년층 감소 등 모집단의 구조가 매우 많이 변화되었으며, 또한 추출틀로 사용한 2005년 인구주택총조사와 가

〈표 2-7〉 지역별 98표본 원표본가구의 유지율

(단위 : %, 가구)

	원표본 수	2차	3차	4차	5차	6차	7차	8차	9차	10차	11차	12차	13차	14차	15차	16차	17차	표본 수
전체	5,000	87.6	80.9	77.3	76	77.2	77.3	76.4	76.4	75.5	74.2	73.1	72.1	70.6	70.3	69.4	69.0	3,451
서울	1,353	84.9	77.6	72.4	70.7	74	72.9	72.4	72	71.8	71.4	69.5	69	66.7	66.5	65.9	65.7	889
부산	486	90.5	89.7	88.3	84.4	80.2	78.4	77.8	77.2	77.8	76.1	74.7	72.6	71.6	72.2	71.2	71.4	347
대구	314	84.1	72	67.5	71	77.4	70.1	73.6	72.3	71.3	68.2	69.4	66.6	65.3	66.9	65.9	65.9	207
대전	153	83.7	81.1	79.7	76.5	80.4	79.7	75.8	73.9	69.9	69.3	70.6	68.6	66	64.7	66.0	64.1	98
인천	298	85.6	81.9	73.2	77.9	75.8	79.5	78.2	79.9	75.8	74.8	75.2	77.2	75.5	74.5	73.5	73.8	220
광주	158	78.5	69	77.2	80.4	79.1	77.2	76.6	75.3	73.4	74.7	70.9	74.1	68.4	67.7	67.1	66.5	105
울산	116	96.6	96.6	89.7	86.2	83.6	82.8	79.3	79.3	74.1	72.4	74.1	75	72.4	72.4	70.7	70.7	82
경기	856	90.2	80	72.7	70.2	75.5	79.7	78.3	77.6	77.3	74.8	73.9	71.6	71.9	70.7	69.7	68.5	586
강원	130	87.7	86.2	73.8	65.4	63.8	67.7	71.5	73.8	75.4	75.4	72.3	70	71.5	70	65.4	66.9	87
충북	113	87.6	75.2	77	75.2	70.8	69.9	67.3	70.8	68.1	69	68.1	70.8	68.1	70.8	71.7	70.8	80
충남	131	91.6	87.8	84	83.2	80.9	80.9	79.4	79.4	81.7	80.2	79.4	77.1	76.3	75.6	77.1	75.6	99
전북	199	88.9	74.4	84.9	83.4	81.9	81.9	80.4	81.9	82.4	81.9	79.4	77.4	74.4	74.4	72.4	71.9	143
전남	127	92.1	89.8	89.8	89	84.3	89	86.6	87.4	88.2	86.6	85.8	84.3	83.5	81.9	80.3	80.3	102
경북	274	89.4	82.5	80.3	82.1	83.6	84.3	82.1	83.2	77.7	76.3	75.9	74.8	72.3	72.3	71.2	69.3	190
경남	292	89.7	88.4	89.4	84.9	83.2	81.2	79.8	80.8	79.8	77.1	76	75.7	74.3	75	73.3	74.0	216

장 최신의 2015년 인구주택총조사의 조사구가 변화되었기 때문에 현재 관리되는 현행 표본이 대표성을 확보하기 위해 확대표본의 검토가 필요한 상황이다.

다. KLIPS에서의 표본 확대 및 패널 관리 방안

'나'항에서 검토한 바와 같이 현행 노동패널은 모집단 변화 및 표본 구조의 변동, 특정 계층의 분석을 위한 표본을 확보하면서 현행 표본이 변화된 모집단의 대표성을 확보하기 위해 표본의 추가가 필요한 상황이다. 현재 논의되고 있는 표본 확대 규모는 5,000가구이다. 그러나 표본의 확대를 위해서 노동패널에서 어떤 목적으로 표본을 추가할 것인가와 어떻게 패널을 관리하는가는 표본 확대 및 관리 방안을 결정하는 중요한 요인이 된다.

제3절의 외국사례에서 살펴본 바와 같이 패널조사에서 표본의 확대 및 관리 방안을 살펴보면, 지역 확대와 같이 변화된 모집단의 커버리지를 확대하는 방안, 이민자 및 설계 시점 이후 국내 이주자를 대상에 포함하는 새로운 대상만을 추가하는 방안, 표본 마모로 인해 변화된 표본 구조의 대표성 유지 및 변화된 모집단에 대한 대표성 확보를 위해 현재 모집단에서 기존 설계틀을 유지하면서 표본을 추가하는 방안, 영국 BHPS나 독일의 SOEP의 혁신패널과 같이 특정 목적을 위한 표본을 추가하는 방안 등을 고려할 수 있다.

한국노동패널에서 표본을 추가하거나 관리하는 방안을 검토하면 다음과 같다.

1) 표본 확대 방안

가) 모집단 대표성의 확보를 위한 표본 추가 방안

모집단 변화 및 표본 구조 변동으로 인해 결여된 모집단 대표성을 확보하기 위해 표본을 추가하는 방안은 추가된 표본을 포함하는 광역시도의 읍면부별 목표 표본크기를 새롭게 결정하여 표본을 추가하는 방안이다.

모집단 대표성을 확보하기 위해 표본을 추가하는 방안은 09통합패널과

같은 방식의 추가 방안으로 2016년 기준에 조사 완료된 통합패널의 원가구와 분가가구수에 추가되는 5,000가구를 더해 현행 모집단 특성을 기준으로 각 층(광역시도의 동읍면부)마다 목표 가구수(전체 약 11,500가구 = 09통합패널 약 6,500가구 + 추가 5,000가구)를 결정한 후 층마다 '목표표본수-현행유지표본수'를 계산하여 부족한 크기만큼 표본을 추가하는 방식이다.

이 방안은 현행 09통합패널을 구축할 시점의 2005년 기준 모집단과 현행 2015년 기준 모집단의 변화를 반영하고, 또한 7차 시점까지 패널의 마모로 인해 결여된 모집단의 대표성을 확보하는 것이 장점인 방안이다.

나) 새로운 계층과 특수 목적 패널 구축을 위한 표본 추가 방안

모집단의 변화와 표본 구조 변동을 보완하면서 현재 모집단의 대표성을 확보하기 위해 표본을 추가하는 방안을 보완하는 방안으로는, 추가표본 규모를 모두 모집단 대표성을 확보하는 크기로 배분하는 것이 아니라, 추가표본의 일부분을 특정 목적이나 특정 계층의 패널로 추가로 확보하기 위해, 별도의 패널 구축을 위해 표본을 추가하는 방안이다. 이는 앞서 외국사례에서 살펴본 바와 같이 특정 조사의 수행이나 목적 수행을 위한 실험 목적의 혁신 표본을 추가하거나, 현재 모집단에서 제외한 외국인 가구나 외국인에 대한 자료를 수집하는 외국인/이민자 표본을 추가하는 방안이다.

노동패널은 보통의 일반가구에 대한 대표성은 충분히 확보되어 있지만 고소득층 가구, 저소득층 가구, 외국인 가구(재외동포 및 조선족 가구 포함) 및 다문화 가구, 새터민 가구, 신혼가구, 청년 1인 가구 등 조사가 어려운 가구에 대한 패널은 충분히 확보하지 못한 상태이므로, 이와 같이 특정 계층의 패널 구축을 위해 표본을 배분하는 방안이다. 또한 노동패널을 이용한 추가적인 조사가 요구되는 상황이 증가하여 기존 노동패널을 이용해 실험적인 목적을 수행하기는 어려우므로, 새로운 조사방법 혹은 새로운 문항 개발 등을 위한 부가적인 패널을 구축해 연구자들이 활용하도록 표본을 추가하는 방안도 포함할 수 있다. 새로운 계층과 특수 목적 패널 구축을 위해 표본을 추가한다면 현행 패널과는 별도로 패널을 관리해야 한다.

이 방안은 현행 패널에 포함되지 않고 있는 특정 계층에 대한 자료 제공을 목적으로 특정 계층에 대한 패널을 확보할 수 있으므로, 현행 패널과 비교 가능한 다양한 계층에 대한 자료를 제공함으로써 노동패널의 활용도가 제고되고, 변화되는 가구의 모집단 특성을 모두 혹은 적절히 반영하는 장점이 있다고 생각한다.

2) 추가 패널의 관리 방안

표본 확대 방안에서 검토한 두 가지 방안으로는, 확대한 패널에 대한 운영 및 관리 방안은 새로운 별도의 패널로 관리하는 방안과 기존 패널에 추가한 확대 패널로 관리하는 방안의 검토가 가능하다. 기본적으로 특정 계층 및 특수 목적을 위한 패널은 전국 단위에서 추정이 가능하도록 별도의 패널로 관리하면 되지만, 현재 모집단의 대표성을 보완하도록 추가되는 패널에 대해서는 패널의 관리 및 운영에 대한 논의가 필요하다.

본항에서는 모집단 변화 및 표본 구조 변동을 반영해 모집단 대표성을 확보하기 위해 추가하는 17통합패널의 관리 및 운영 방안을 검토하기로 한다.

가) 새로운 별도의 패널로 관리하는 방안

새로운 독립 패널로 관리하는 방안은 기존 2개의 패널(98패널, 09통합패널)과 구별해 별도로 제3의 패널(17통합패널)로 운영하는 것을 목적으로 관리하는 방안이다.

현재 노동패널은 최초 시점인 1998년에 도시지역 가구만을 대상으로 설계된 원패널(98표본)과 2009년 당시 전국으로 지역을 확대하기 위해 표본을 추가하여 구축한 2009년 통합패널로 별도로 관리되고 있다. 2009년 통합패널은 98패널에 읍면부를 포함하여 전국으로 확대하기 위해 2008년도 조사완료된 가구와 추가된 가구를 더해 광역시도의 동부 및 읍면부로 목표 규모를 결정한 후 기존 98패널크기를 제외한 다음, 광역시도의 동읍면부별로 부족한 표본만을 추가하여 확대하였다. 09통합패널은 원패널과 달리 읍면부로의 커버리지를 확대하고 모집단의 전국적인 가구 대표성을 확보하기 위해 추가한 것이므로 기존의 98패널과 분리하여 별도의 통합패널로 관리하고 있다. 09통합패널은 98패널의 원가구패널과 분가가구패

널 중 2009년 표본 확대 시점에서 조사된 모든 패널과 추가된 패널을 09통합패널의 원패널로 하여 구축함으로써 09통합패널은 98원패널과 구별해 관리하고 있다. 09통합패널의 개인 패널은 2009년 통합 시점에 조사된 모든 원가구원, 신규 유입 원가구원 및 비원가구원을 모두 포함해 통합패널의 원가구원으로 정의하고 있으며, 원가구원 중 2009년 조사에 응답하지 않은 원가구원은 노동패널의 신규 및 유입 규칙에 따라 신규 유입가구원으로 추가하지만 비원가구원은 제외하고 있다. 따라서 현재 노동패널은 98패널과 09통합패널로 관리되고 패널마다 별도의 가중치를 제공하고 있다. 즉 98원패널은 98패널 횡단 및 종단 가중치와 09통합패널 횡단 및 종단 가중치를 별도로 부여하고, 98패널의 비원가구원 패널은 09통합패널에서는 원가구원으로 포함되기에 98패널 횡단 가중치와 09통합패널의 횡단 및 종단 가중치를 부여받고, 09통합패널로 추가된 원가구원은 09통합패널의 횡단 및 종단 가중치만을 부여받는다.

이처럼 2017년에 새롭게 추가되는 패널을 기존의 2개의 패널과 구분해 새로운 별도의 패널로 관리하고 운영하는 방안이다. 노동패널을 일관성 있게 운영하고 관리하기 위해 17통합패널의 운영 및 관리 방식은 09통합패널과 동일하게 적용하면 된다. 그러나 17통합패널은 기존의 98패널, 09통합패널과 구분해 별도의 패널로 가정하므로 2015년 인구주택총조사를 기준으로 17통합패널을 위한 가중치를 산정해 별도로 제공해야 한다.

이 방안은 제3의 노동패널을 제공하는 것으로 현행 2개의 패널과는 별도로 관리해야 하고, 또한 이용자를 위해 3개의 데이터셋으로 구분해 제공해야 하며, 동시에 가중치도 별도로 산정해야 하므로 그 작업의 양이 증가하는 단점이 있다. 특히 98원패널은 98패널 가중치, 09통합패널 가중치 및 17통합패널의 가중치를 각각 산정해야 하고, 09통합 패널은 09통합패널 가중치와 17통합패널의 가중치를 각각 산정해야 한다. 하지만 가중치 산정은 별도의 모집단을 기준으로 부여하므로 가중치는 어렵지 않게 쉽게 산출이 가능하다. 즉 17통합패널은 2015년 인구주택총조사를 기준으로 가중치를 산정하면 된다.

나) 09통합패널을 보완한 확대패널로 관리하는 방안

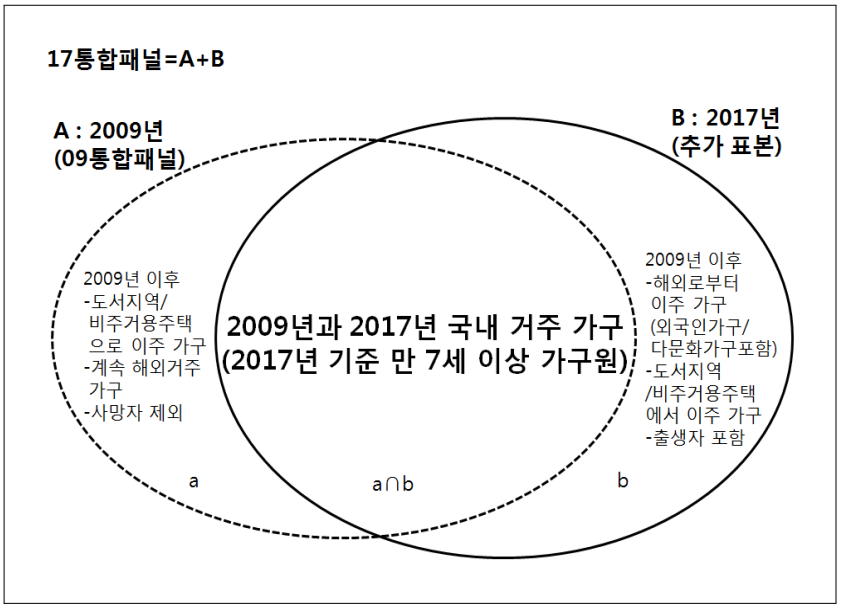
노동패널의 98패널은 도시가구 및 가구원을 대상으로 구축된 패널이지

만 09통합패널과 17통합패널은 전국의 일반가구를 대상으로 하는 패널이다. 09통합패널과 17통합패널을 별도로 관리하지 않고 17통합 패널로 통합 관리하는 방안을 검토해 보고자 한다.

일반적으로 패널이 장기화되면 설계 시점의 모집단과 횡단 시점의 현재 모집단의 특성 차이가 심하게 되고, 또한 패널 무응답 및 패널의 마모가 증가하여 표본 구조가 변화되어 기존 패널이 횡단 모집단의 대표성을 유지하기 어렵게 된다. 특히, 표본의 구조 및 특성이 변화된 경우는 적절한 사후층화 가중치를 부여해 횡단 가중치를 조정하더라도 근본적으로 수집된 표본 특성값이 모집단 특성값과의 차이를 제거하거나 감소시키지 못하므로 추정 및 분석 결과는 바람직하지 못한 편향된 결과를 제공하게 될 것이다. 이를 위해 조사방법론에서는 가중치를 조정하기보다는 표본의 추가 혹은 확대를 하거나 기존 표본을 폐기하고 새로운 표본을 구축하는 것이 일반적인 방법이다. 그러나 패널조사에서는 기존 패널을 폐기하고 새로운 패널을 구축하는 것보다는 표본을 추가하여 패널을 확대하는 방안이 더 효과적인 방안이 된다.

09통합패널을 보완한 확대패널로 관리하는 방안은, [그림 2-6]처럼 17통합패널은 09통합패널의 모집단과 2017년 모집단을 결합한 형태의 모집단에서 각각 추출된 표본으로 패널을 구축하는 것으로 가정한다. 즉 09통합패널에 2017년 표본을 추가한 17통합패널은 2009년 모집단과 2017년 모집단으로부터 각각 표본을 추출하여 결합한 결합표본의 구조로 간주하는 것이다. 2009년을 기준으로 17통합패널의 모집단 구조를 살펴보면, 2009년과 2017년에 모두 존재하는 일반가구 및 가구원($a \cap b$)과 2009년 시점에서 제외된 모집단에서 2017년에 일반가구 및 가구원으로 추가된 부분(b)을 포함하지만 2009년 모집단에서 자연적으로 제외된 부분(a)은 제외한다. 여기서 2017년에 추가된 부분(b)에 09년 이후 신규 혹은 분가가구 및 가구원을 제외하는 이유는 09통합패널의 분가가구 및 가구원으로 현행 패널에 반영되어 있기 때문이다. 그러므로 2017년 모집단에서는 2009년 당시 제외된 가구 및 가구원 중 2017년에 일반가구로 추가된 부분에 대해서만 패널을 추가한 결합 표본 구조로 17통합패널을 구축한다는 것이다.

〔그림 2-6〕 KLIPS의 결합표본의 결합 모집단 구조와 유형



09통합패널을 확대패널로 관리하는 방안은, 98원표본은 2종의 통합패널과는 성격이 다르므로 별도로 제공하도록 관리하는 것이 가능하지만 2종의 통합패널(09패널, 17패널)은 전국의 일반가구를 대상으로 하는 패널이므로 통합한 확대패널(혹은 결합패널)은 09통합패널은 제공하지 않고 17통합패널(혹은 17확대패널)만을 제공하도록 관리하는 방안이다. 17통합패널을 제공한다면, 앞서 설명한 바와 같이 새로운 별도의 패널로 관리하여 3종의 노동패널 자료를 제공한다면 시간 및 비용이 증가할 뿐만 아니라 이용자 관점에서 횡단분석을 위해 어느 패널자료를 이용해 분석해야 하는가 하는 혼란을 예방할 수 있으며, 현재의 횡단 모집단의 대표성을 유지하므로 추정 및 분석 결과의 정도가 높아질 것이다.

라. 소 결

본항에서는 2017년에 추가표본을 추출하여 패널을 확대한다는 가정에서 표본 확대 방안과 패널의 관리 방안을 검토하였다. 검토한 주요한 결

과를 간단히 요약하면 다음과 같다.

먼저, 표본 추가에 대한 필요성 및 확대 방안을 검토해 보면, 일부분의 표본으로 배분해 특정 계층을 위한 별도의 패널을 구축하는 것도 필요하지만, 본 연구에서는 KLIPS의 추가표본 방안으로 표본 이탈 및 표본 마모로 인해 변화된 표본구조의 대표성을 유지하고 변화된 모집단의 대표성을 확보하면서 지역 및 특정 계층의 통계 생산이 가능하고, 현재 모집단에서 기존 설계틀을 유지하도록 표본을 추가하여 패널을 확대하는 방안이 바람직하다고 생각한다.

둘째, 보완적으로, 직관적으로 대표성 유지를 위해 표본을 확대하는 동시에 변화된 모집단 특성을 반영하고 현재 패널에서 관리가 어려운 특정 계층 혹은 특수 목적을 위한 별도의 패널을 추가하는 방안의 검토가 필요하다고 생각한다. 2010년과 비교한 2015년 모집단은 1인 가구의 증가, 인구 감소로 인한 청년층의 감소, 도시 및 남성 1인 가구의 증가, 다문화 가구의 증가 등 다양한 특성의 변화가 예상되기 때문에 다양한 모집단 특성에 대한 노동 및 고용 관련 자료를 제공할 필요가 있으므로 이에 대한 표본 추가 방안의 검토가 요구된다.

셋째, 기존 패널을 추가한 통합패널의 관리 방안을 검토하여 보면, 본 연구에서 제안하는 것처럼 추가표본을 원표본에 결합해 확대표본(the expanded sample)을 운영하는 방안은 기존 패널과는 별도로 현재 시점의 모집단에서 새로운 표본을 추출하여 독립된 3개의 패널(98원패널, 09통합패널, 17통합패널)로 운영하는 방안과, 09통합패널의 1차 조사 이후 변화된 모집단 특성을 반영하기 위해 현재 모집단에서 표본을 추가 추출해 현행 09통합패널과 결합해 확대패널을 구축함으로써 현행과 같이 2개의 패널(98패널, 17확대패널)로 관리하는 방안의 검토가 가능하다. KLIPS의 경우, 현실적으로 독립적인 3개의 패널을 운영하는 것은 패널 관리를 위한 예산 및 인력, 시간 등이 많이 필요하고, 또한 표본 확대의 필요성을 고려한다면 변화된 모집단에서 추출된 표본을 기존 표본에 추가해 결합표본(the combine sample)을 운영하는 것이 효과적이다. 이를 위해서는 09통합패널은 09통합패널의 2009년 기준의 종단 분석을 위한 자료로만 제공하고, 2017년 이후는 횡단 및 종단 분석을 위해서는 17확대패널만을

제공하도록 관리하는 방안이 바람직하다.

결론적으로, 본 연구에서는 [그림 2-6]과 같이 KLIPS의 98원표본과 09통합표본의 원패널에 현재시점(2016년 기준)의 모집단에서 추출된 표본을 추가한 결합표본으로 표본을 확대하는 방식으로 관리하는 방안을 제안한다. 제안된 방안을 고려한다면, 98패널과 09통합패널도 17확대패널과는 구분해 별도로 관리하지만 노동패널자료는 98패널과 17확대패널자료만을 제공하는 방안이 바람직하다고 생각한다. 이를 위해 기존의 98패널은 종단 분석을 위해 종단 가중치를 부여해 자료를 제공해야 하고, 17확대패널은 횡단 가중치 및 종단 가중치를 부여해 제공하는 방안의 검토가 필요하다. 필요하다면 09통합패널은 일반 공개보다는 종단 연구를 위해 별도로 제공하는 방안의 검토가 가능할 것이다. 참고로, 17통합패널은 결합 표본 구조이므로 모집단 구조를 반영한 가중치 산출과 추정 과정에 대한 논의가 추가적으로 필요하다. 결합 구조의 패널에 대한 가중치 산출 및 추정 과정에 대해서는 2항 나.에서 호주의 HILDA의 가중치 산출 및 추정 과정을 소개하기로 한다.

2. 표본 확대에 따른 KLIPS 결합표본의 모집단과 가중치

위항에서 살펴본 바와 같이 KLIPS는 표본 이탈 및 표본 마모, 표본의 고령화, 지역통계 생산 및 특정 계층 분석, 변화된 모집단 구조를 반영하면서 종단 및 횡단 모집단의 대표성 유지 및 확보를 위해 표본의 확대가 필요하다.

본항에서는 KLIPS의 98원표본과 09통합표본의 원패널에 현재시점의 모집단에서 추출된 표본을 추가하는 방식으로 표본을 확대해 결합표본을 구성하는 방안을 제안하기 위해 KLIPS의 추가표본 추출 및 결합표본에 대한 모집단 및 추출틀의 작성 방안과, 결합표본의 추정 및 분석을 위한 가중치 산출 방법과 과정에 대한 방안을 검토해 본다.

가. KLIPS 결합표본의 모집단과 표본추출틀

1) 2010년 인구주택총조사의 모집단과 추출틀

가) 2010년 인구주택총조사의 모집단과 추출틀

통계청에서는 매 5년 단위로 조사원이 가구를 방문하는 현장조사 기반의 인구주택총조사를 실시하여 상주인구 기준의 가구 및 인구 모집단을 설정하고, 주택유형을 고려해 표본 추출을 위한 평균 60가구 기준의 조사구를 생성한다. 본 연구에서는 2010년 인구주택총조사의 추출틀을 사용하

〈표 2-8〉 2010년 인구주택총조사의 모집단과 추출틀의 포함 범위

	2010년인구주택총조사			2010 총조사 추출틀					
				소계		일반조사구		아파트조사구	
	가구수	인구수	가구당 인구	조사구수	가구수	조사구수	가구수	조사구수	가구수
전국	17,499,195	46,436,685	2.65	303,180	17,499,195	169,214	9,467,751	133,966	8,031,444
서울	3,574,936	9,429,798	2.64	0,925	3,574,936	37,779	2,173,518	23,146	1,401,418
부산	1,249,819	3,299,836	2.64	21,716	1,249,819	11,763	654,765	9,953	595,054
대구	873,557	2,374,112	2.72	15,211	873,557	7,926	437,251	7,285	436,306
인천	920,471	2,560,167	2.78	15,176	920,471	7,513	460,114	7,663	460,357
광주	518,276	1,421,291	2.74	9,010	518,276	3,653	191,968	5,357	326,308
대전	535,947	1,439,526	2.69	8,957	535,947	4,301	254,697	4,656	281,250
울산	377,523	1,040,506	2.76	6,519	377,523	3,235	182,654	3,284	194,869
세종	45,095	108,496	2.41	790	45,095	447	24,079	343	21,016
경기	3,905,363	10,897,852	2.79	66,578	3,905,363	32,222	1,826,846	34,356	2,078,517
강원	560,024	1,390,197	2.48	10,179	560,024	5,955	318,915	4,224	241,109
충북	564,019	1,437,355	2.55	10,059	564,019	5,783	311,174	4,276	252,845
충남	756,381	1,897,648	2.51	13,560	756,381	8,228	445,401	5,332	310,980
전북	661,787	1,693,848	2.56	11,811	661,787	6,749	359,804	5,062	301,983
전남	641,290	1,579,118	2.46	11,618	641,290	7,649	404,748	3,969	236,542
경북	1,013,197	2,456,156	2.42	18,248	1,013,197	12,140	655,905	6,108	357,292
경남	1,160,110	3,009,319	2.59	20,336	1,160,110	11,666	640,711	8,670	519,399
제주	186,495	509,956	2.73	3,277	186,495	2,652	149,280	625	37,215

주 : 2010년 총조사 시점의 주소를 이용해 세종시의 가구 및 인구수를 추계하여 산출함.

〈표 2-9〉 세종시 센서스 결과(2013년, 통계청)

행정구역	거처수 (호)	일반가구(외국인가구 제외)	
		가구수	가구원수
세종	36,905	44,523	108,706
동부	7,011	7,001	21,940
읍부	13,983	20,052	47,031
면부	15,911	17,470	39,735

는 방안을 검토하기 위해 2010년 인구주택총조사 모집단을 살펴본다. 2010년 인구주택총조사에서 외국인가구와 외국인을 제외한 일반 모집단과 추출틀에 포함(coverage)된 모집단 규모는 <표 2-8>, <표 2-9>와 같다.

<표 2-8>을 보면, 2010년 인구주택총조사의 일반 모집단은 가구가 1,750만 가구이고, 인구는 4,644만 명으로 가구당 평균 2.65명이다. 표본틀로 사용하는 2010년 인구주택총조사의 조사구는 모두 303,180개이다. 단독/연립/다세대의 일반조사구는 169,214개이고, 가구는 약 947만 가구이며, 아파트조사구는 133,966개이고, 가구는 약 803만 가구이다. 세종특별자치시는 2013년에 충남과 충북에서 분리되므로 2010년 총조사에서는 충남과 충북에 포함된다. <표 2-8>에서 세종시는 2013년 세종시로 편입된 충남과 충북지역에서 2010년의 주소로 추산한 인구이고, <표 2-9>는 2013년 세종시로 분리된 이후 실시한 세종시센서스 결과이다.

<표 2-10>은 단독/연립/다세대의 일반조사구 추출틀에 대한 조사구 특성자료이다. 조사구당 가구수는 평균 56가구이고, 인구수는 평균 130명이다. <표 2-11>은 아파트조사구 추출틀에 대한 조사구 특성자료이다. 조사구당 가구수는 평균 60가구이고, 인구수는 평균 182명이다. 아파트조사구가 일반조사구보다 평균 가구수 및 인구수가 많으며, 지역에 따라 평균 가구수 및 인구수의 변동이 있다.

대부분의 표본조사에서 사용하는 기본적인 표본추출틀로는 인구주택총조사의 조사구를 널리 사용한다. 그러나 표본설계 시점이 추출틀의 작성 시점과 설계 시점의 차이가 커지면 인구주택총조사 모집단의 추출틀은 현재 모집단을 적절하게 표현하지 못한다. 2010년 총조사 이후 시간이

〈표 2-10〉 2010년 인구주택총조사의 일반조사구 추출틀

	가구수	조사구당 평균가구수	조사구수	인구수	조사구당 평균인구수
전국	9,467,751	55.95	169,214	22,030,286	130.19
서울	2,173,518	57.53	37,779	5,157,853	136.53
부산	654,765	55.66	11,763	1,560,243	132.64
대구	437,251	55.17	7,926	1,051,890	132.71
인천	460,114	61.24	7,513	1,119,626	149.03
광주	191,968	52.55	3,653	432,159	118.30
대전	254,697	59.22	4,301	580,085	134.87
울산	182,654	56.46	3,235	433,322	133.95
세종	24,079	53.87	447	47,548	106.37
경기	1,826,846	56.70	32,222	4,337,788	134.62
강원	318,915	53.55	5,955	718,944	120.73
충북	311,174	53.81	5,783	705,938	122.07
충남	445,401	54.13	8,228	990,699	120.41
전북	359,804	53.31	6,749	807,013	119.58
전남	404,748	52.92	7,649	888,173	116.12
경북	655,905	54.03	12,140	1,412,238	116.33
경남	640,711	54.92	11,666	1,436,327	123.12
제주	149,280	56.29	2,652	397,988	150.07

〈표 2-11〉 2010년 인구주택총조사의 아파트조사구 추출틀

	가구수	조사구당 평균가구수	조사구수	인구수	조사구당 평균인구수
전국	8,031,444	59.95	133,966	24,406,399	182.18
서울	1,401,418	60.55	23,146	4,271,945	184.57
부산	595,054	59.79	9,953	1,739,593	174.78
대구	436,306	59.89	7,285	1,322,222	181.50
인천	460,357	60.08	7,663	1,440,541	187.99
광주	326,308	60.91	5,357	989,132	184.64
대전	281,250	60.41	4,656	859,441	184.59
울산	194,869	59.34	3,284	607,184	184.89
세종	21,016	61.27	343	60,948	177.69
경기	2,078,517	60.50	34,356	6,560,064	190.94
강원	241,109	57.08	4,224	671,253	158.91
충북	252,845	59.13	4,276	731,417	171.05
충남	310,980	58.32	5,332	906,949	170.10
전북	301,983	59.66	5,062	886,835	175.19
전남	236,542	59.60	3,969	690,945	174.09
경북	357,292	58.50	6,108	1,043,918	170.91
경남	519,399	59.91	8,670	1,572,992	181.43
제주	37,215	59.54	625	111,968	179.15

지나감에 따라 모집단은 지역별, 주택유형별 등 가구, 주택 및 인구학적 특성에 따라 변화하지만 매 5년 주기로 인구주택총조사를 실시하므로 2010년 이후 다음 인구주택총조사 이전에 수행되는 표본설계 시점에서는 현재 시점의 표본틀에 대한 정확한 정보를 알 수 없다. 이처럼 추출틀의 작성 시점과 설계 시점에 시점 차이가 발생하는 경우 모집단은 설계 시점의 모집단으로 정의하고, 표본추출을 위한 추출틀로는 ① 인구주택총조사 조사구에 최신 벤치마크 자료를 결합하여 보완한 인구주택총조사의 현행 조사구를 이용하거나, ② 주민등록자료로 모집단을 설정하여 통-반-리를 조사구로 활용하는 방안 중 하나를 선택한다. 그러나 주민등록인구자료는 인구에 대한 다양한 현재 모집단 상황을 보여주지만 행정등록 기준의 자료이고, 조사구를 대체하는 추출틀로 사용하는 통-반-리의 세대 크기가 지역에 따라 세대수 및 인구수가 일정하지 않은 단점이 있어 인구주택총조사의 조사구를 추출틀로 사용하되 벤치마크 자료를 이용해 보완한 추출틀을 사용한다. 즉 대부분의 표본조사에서는 현재 모집단의 가구와 인구에 대한 모집단 자료로 추계자료를 정의하지만 추계자료는 지역별, 동읍면부별, 성별 및 연령대별 등의 단위에 대한 모집단만을 제공하고 추출틀에 대한 정보를 제공하지 않기 때문에 설계 시점에서 최신 인구주택총조사의 조사구를 벤치마크 자료로 보완한 추출틀을 널리 사용한다. 추출틀을 보완하기 위한 모집단 주택유형의 벤치마크 자료로는 전국의 모든 주택에 대한 가격동향을 파악하는 기초자료를 제공하는 주택가격 동향자료(한국감정원)를 널리 이용한다. 전국의 개별 주택 및 단지별로 주택유형, 가격, 면적 등의 자료가 포함되어 있다.

나) 신축아파트를 추가한 모집단의 추출틀

KLIPS의 추가표본 추출 및 결합표본의 모집단은 2016년도 모집단에 속하는 가구와 15세 이상 개인이다. 추가표본 및 결합표본의 현재 모집단은 추계자료(추계 가구, 추계 인구)를 이용하지만 추계자료는 현재 모집단의 추계된 가구수 및 인구수만을 제공할 뿐 추출틀(조사구)의 정보를 제공하지 않는다. 최근에 수행된 인구주택총조사를 보면, 2010년 이전은 현장조사 기반으로, 2015년부터는 행정자료 기반의 등록센서스로 실시된

다. 최신 수행된 2015년 등록센서스 기반의 추출틀은 2017년 상반기(3,4월경)에 공개될 예정이므로 현시점에서는 2010년 인구주택총조사를 추가표본의 추출틀로 사용할 수밖에 없는 상황이다. 추가표본의 추출틀인 인구주택총조사 조사구는 매 5년마다 작성되기 때문에 추출틀을 사용하는 시점과 추출틀의 작성 시점의 차이가 클수록 변화된 모집단 특성을 왜곡하게 되므로 최신의 인구주택총조사 조사구 및 최신 모집단 자료를 벤치마크 자료로 결합하여 보완하거나 수정하여 추출틀(조사구)을 사용하는 것이 바람직하다.

본항에서는 추계 자료를 모집단으로 정의하고, 추가표본의 추출틀을 작성하기 위해 2010년 인구주택총조사의 조사구와 벤치마크 자료인 2010년 총조사 이후 신축아파트 단지를 결합한 추출틀을 검토한다.

(1) 신축아파트 현황

<표 2-12>는 2010년 총조사 이후인 2011년 1월 1일부터 2015년 상반기(7월)까지의 신축아파트 현황자료이다. 2010년 총조사 이후 모두 약 134만 세대의 아파트가 신축되었고, 시지역에 124만 세대, 군지역에 약 10만 세대가 공급된 것이다. 2010년 총조사 이후 특정시점까지의 신축아파트 세대수를 반영하면 총조사 이후 특정시점까지의 변화된 모집단 특성과 구조를 파악하는 데 유용하다. 그러나 일반주택에 대한 벤치마크 자료로 이용되는 유용한 자료가 없기 때문에 총조사 이후 신축아파트만 반영하여 변화된 모집단의 구조와 특성을 반영하게 된다. <표 2-12>를 보면, 경기, 서울, 부산, 인천 등 수도권 및 대도시지역, 그리고 읍면지역인 군부보다는 시부에서 모집단 변화가 많은 것으로 나타난다.

(2) 신축아파트를 추가한 추출틀

<표 2-13>은 추가표본 및 결합표본의 추출틀로 사용하기 위해 2010년 인구주택총조사의 조사구와 총조사 이후 신축아파트세대수 자료를 벤치마크하여 수정한 추출틀 자료이다. 2010년 총조사에서는 약 1,750만 가구였으나 신축아파트를 반영하면 가구수는 약 1,884만 가구로 증가된다. 최신의 신축아파트 자료를 반영함으로써 2010년 총조사자료만을 사용하는 한계를 보완해 주지만, 일반주택에 대한 벤치마크 자료가 없기 때문에 변

화된 모집단의 구조와 특성을 충분히 설명하지 못한다는 점을 유념해 사용해야 한다.

2) 2015년 인구주택총조사의 등록센서스 기반 모집단과 추출틀

가) 2015년 인구주택총조사의 모집단 특성

2015년 인구주택총조사는 기존의 현장조사 기반 인구주택총조사와는 달리 각 정부부처의 행정자료를 기반으로 하는 등록 자료 기반의 등록센

〈표 2-12〉 신축아파트 추출틀(2010년 인구주택총조사 이후 신축아파트)

	단지수			세대수		
	전체	시부	군부	전체	시부	군부
전국	2,665	2,047	618	1,338,180	1,235,400	102,780
서울	284	284	0	256,158	256,158	0
부산	143	120	23	122,766	105,074	17,692
대구	90	78	12	35,798	32,049	3,749
인천	86	70	16	102,250	100,954	1,296
광주	63	63	0	39,074	39,074	0
대전	67	67	0	39,579	39,579	0
울산	58	38	20	30,967	26,456	4,511
세종	19	19	0	27,626	27,626	0
경기	616	557	59	363,553	358,402	5,151
강원	100	55	45	22,077	18,392	3,685
충북	135	82	53	30,798	20,192	10,606
충남	179	97	82	50,422	37,052	13,370
전북	122	88	34	37,916	32,374	5,542
전남	146	61	85	43,752	32,701	11,051
경북	240	168	72	42,331	37,470	4,861
경남	243	126	117	71,297	50,031	21,266
제주	74	74	0	21,816	21,816	0

주 : 혼합평형 단지는 평형별 별도로 간주해 산출한 수치이므로 실제 단지수와는 다름.

자료 : 2011년 1월 1일 이후부터 2015년 상반기(7월)까지의 신축아파트 자료(한국감정원).

〈표 2-13〉 신축아파트를 결합한 2010년 인구주택총조사 기반의 추출틀 특성

	가구수 (+세대수)	2010년 인구주택총조사						신축 아파트
		소계		일반조사구		아파트조사구		
		조사구수	가구수	조사구수	가구수	조사구수	가구수	세대수
전국	18,837,375	303,180	17,499,195	169,214	9,467,751	133,966	8,031,444	1,338,180
서울	3,831,094	60,925	3,574,936	37,779	2,173,518	23,146	1,401,418	256,158
부산	1,372,585	21,716	1,249,819	11,763	654,765	9,953	595,054	122,766
대구	909,355	15,211	873,557	7,926	437,251	7,285	436,306	35,798
인천	1,022,721	15,176	920,471	7,513	460,114	7,663	460,357	102,250
광주	557,350	9,010	518,276	3,653	191,968	5,357	326,308	39,074
대전	575,526	8,957	535,947	4,301	254,697	4,656	281,250	39,579
울산	408,490	6,519	377,523	3,235	182,654	3,284	194,869	30,967
세종	72,721	790	45,095	447	24,079	343	21,016	27,626
경기	4,268,916	66,578	3,905,363	32,222	1,826,846	34,356	2,078,517	363,553
강원	582,101	10,179	560,024	5,955	318,915	4,224	241,109	22,077
충북	594,817	10,059	564,019	5,783	311,174	4,276	252,845	30,798
충남	806,803	13,560	756,381	8,228	445,401	5,332	310,980	50,422
전북	699,703	11,811	661,787	6,749	359,804	5,062	301,983	37,916
전남	685,042	11,618	641,290	7,649	404,748	3,969	236,542	43,752
경북	1,055,528	18,248	1,013,197	12,140	655,905	6,108	357,292	42,331
경남	1,231,407	20,336	1,160,110	11,666	640,711	8,670	519,399	71,297
제주	208,311	3,277	186,495	2,652	149,280	625	37,215	21,816

서스로 2015년 11월 1일을 기준으로 실시하였다. 2016년 9월 공개된 2015 등록센서스의 모집단 자료는 <표 2-14>와 같다(자료 : 통계청, 「2015 등록센서스」, 2016.9.).

<표 2-14>를 보면, 2015년도 우리나라 전체 가구는 1,956만 가구이고, 인구(개인)는 5,107만 명이다. 모집단을 보면, 가구 및 인구의 약 81%가 동부(시부)에 거주하고 있고, 수도권에 49~49.5%, 특광역시 등 대도시에 약 45~46%, 특히 경기도에 23~24.5%가 거주하고 있다.

<표 2-15>의 가구 현황을 보면, 우리나라 전체 가구 중 집단 및 외국인

가구를 제외한 일반가구가 1,911만 가구(전체 가구의 97.7%)이다. 특히 일반가구 중 1인 가구가 520만 가구(27.2%), 2인 가구가 499만 가구(26.1%)로 일반가구 중 2인 이하 가구가 전체의 53.3%나 되어 2010년 인구주택총조사에 비해 2인 이하 가구가 급격히 증가하였음을 보여주고 있다.

〈표 2-14〉 2015년 인구주택총조사의 등록센서스 결과

행정구역	인 구	구성비	가 구	구성비
전 국	51,069,375	100.0	19,560,603	100.0
읍 부	4,616,802	9.0	1,737,415	8.9
면 부	4,774,878	9.3	1,981,033	10.1
동 부	41,677,695	81.6	15,842,155	81.0
수도권	25,273,824	49.5	9,518,698	48.7
특광역시	23,121,530	45.3	8,939,058	45.7
서울특별시	9,904,312	19.4	3,914,820	20.0
부산광역시	3,448,737	6.8	1,348,315	6.9
대구광역시	2,466,052	4.8	937,573	4.8
인천광역시	2,890,451	5.7	1,066,297	5.5
광주광역시	1,502,881	2.9	573,181	2.9
대전광역시	1,538,394	3.0	588,395	3.0
울산광역시	1,166,615	2.3	434,058	2.2
세종특별자치시	204,088	0.4	76,419	0.4
도지역	27,947,845	54.7	10,621,545	54.3
경기도	12,479,061	24.4	4,537,581	23.2
강원도	1,518,040	3.0	611,578	3.1
충청북도	1,589,347	3.1	613,004	3.1
충청남도	2,107,802	4.1	816,247	4.2
전라북도	1,834,114	3.6	726,572	3.7
전라남도	1,799,044	3.5	730,743	3.7
경상북도	2,680,294	5.2	1,078,479	5.5
경상남도	3,334,524	6.5	1,282,617	6.6
제주특별자치도	605,619	1.2	224,724	1.1

〈표 2-15〉 2015년 인구주택총조사 가구 현황

행정구역	가구	일반가구						집단	외국인
			1인	2인	3인	4인	5인 이상		
전 국	19,560,603	19,111,030	5,203,440	4,993,818	4,100,979	3,588,931	1,223,862	16,464	433,109
읍 부	1,737,415	1,695,508	448,466	475,427	350,875	299,401	121,339	2,189	39,718
면 부	1,981,033	1,927,621	630,013	673,690	319,506	196,819	107,593	4,284	49,128
동 부	15,842,155	15,487,901	4,124,961	3,844,701	3,430,598	3,092,711	994,930	9,991	344,263
수도권	9,518,698	9,214,649	2,385,893	2,257,768	2,067,905	1,893,629	609,454	6,853	297,196
특광역시	8,939,058	8,742,627	2,419,096	2,201,794	1,921,350	1,671,734	528,653	5,396	191,035
서울	3,914,820	3,784,490	1,115,744	930,467	817,440	701,945	218,894	2,261	128,069
부산	1,348,315	1,335,900	361,749	366,048	296,742	238,031	73,330	686	11,729
대구	937,573	928,528	239,517	239,824	208,795	184,258	56,134	574	8,471
인천	1,066,297	1,045,417	243,678	265,079	245,135	220,538	70,987	713	20,167
광주	573,181	567,157	163,577	137,662	115,701	109,612	40,605	438	5,586
대전	588,395	582,504	169,391	140,603	122,088	112,055	38,367	420	5,471
울산	434,058	423,412	103,551	104,101	100,230	90,735	24,795	239	10,407
세종	76,419	75,219	21,889	18,010	15,219	14,560	5,541	65	1,135
도지역	10,621,545	10,368,403	2,784,344	2,792,024	2,179,629	1,917,197	695,209	11,068	242,074
경기	4,537,581	4,384,742	1,026,471	1,062,222	1,005,330	971,146	319,573	3,879	148,960
강원	611,578	606,117	189,379	178,470	116,384	87,182	34,702	788	4,673
충북	613,004	601,856	173,598	167,146	120,634	100,527	39,951	814	10,334
충남	816,247	796,185	234,513	226,658	153,585	128,382	53,047	1,107	18,955
전북	726,572	717,311	213,750	209,008	135,856	110,132	48,565	834	8,427
전남	730,743	720,612	218,864	227,975	131,454	96,940	45,379	933	9,198
경북	1,078,479	1,062,724	322,569	320,498	207,395	156,364	55,898	1,312	14,443
경남	1,282,617	1,258,487	346,754	341,745	265,127	228,216	76,645	1,176	22,954
제주	224,724	220,369	58,446	58,302	43,864	38,308	21,449	225	4,130

특히 2인 이하 가구는 과거와 동일하게 읍면지역에 고령화 가구로 인해 높게 나타났지만 2010년에 비해 동부 및 수도권 지역에 많은 증가를 보이고 있다.

<표 2-16>의 인구 현황을 보면, 우리나라 전체 내국인 인구의 성비는 큰 차이가 없이 거의 동일하지만, 생산연령인구는 남성이 약간 높고, 고령인구는 여성이 20% 정도 더 많다. 특이한 점은 고령인구 및 생산연령인구는 도시지역보다 읍면지역에서 남성의 비율이 더 높다. 2010년 인구

〈표 2-16〉 2015년 인구주택총조사 인구 현황

행정구역	인구(내국인)		생산연령인구(15~64세)		고령인구(65세 이상)	
		남자		남자		남자
전 국	49,705,663	24,819,839	36,229,874	18,489,317	6,569,082	2,763,407
읍 부	4,467,697	2,262,853	3,105,454	1,629,374	659,536	270,140
면 부	4,546,520	2,291,860	2,848,845	1,555,808	1,270,881	513,589
동 부	40,691,446	20,265,126	30,275,575	15,304,135	4,638,665	1,979,678
수도권	24,416,226	12,148,910	18,201,546	9,182,336	2,794,103	1,206,311
특광역시	22,568,209	11,183,692	16,851,257	8,466,147	2,770,294	1,196,530
서울	9,567,196	4,694,317	7,210,830	3,569,884	1,202,894	530,871
부산	3,404,667	1,675,339	2,503,218	1,253,692	499,586	214,040
대구	2,436,770	1,211,219	1,799,783	911,892	311,127	129,779
인천	2,822,601	1,414,793	2,113,068	1,077,497	305,511	129,507
광주	1,481,289	736,656	1,089,468	549,632	165,765	69,904
대전	1,519,314	763,310	1,132,777	578,144	163,782	70,138
울산	1,136,755	587,603	863,083	453,824	100,641	43,601
세종	199,617	100,455	139,030	71,582	20,988	8,690
도지역	27,137,454	13,636,147	19,378,617	10,023,170	3,798,788	1,566,877
경기	12,026,429	6,039,800	8,877,648	4,534,955	1,285,698	545,933
강원	1,499,734	758,601	1,052,334	552,869	254,088	105,744
충북	1,548,589	779,961	1,102,597	573,321	229,445	94,768
충남	2,036,720	1,034,139	1,411,048	743,934	331,110	137,883
전북	1,804,184	899,308	1,234,821	640,791	323,552	131,542
전남	1,764,433	879,860	1,155,961	609,717	371,685	148,122
경북	2,622,729	1,314,196	1,818,751	949,578	467,953	189,348
경남	3,244,163	1,634,407	2,314,293	1,205,772	452,232	180,075
제주	590,473	295,875	411,164	212,233	83,025	33,462

〈표 2-17〉 2010년과 2015년 모집단 비교

	2010년 인구주택총조사			2016년 등록센서스			2016년 등록센서스 (일반가구/내국인)		
	가구수	인구수	가구당 인구	가구수	인구수	가구당 인구	일반가구 가구수	내국인 인구수	가구당 인구
전국	17,499,195	46,436,685	2.65	19,560,603	50,617,045	2.59	19,111,030	49,705,663	2.60
서울	3,574,936	9,429,798	2.64	3,914,820	9,904,312	2.53	3,784,490	9,567,196	2.53
부산	1,249,819	3,299,836	2.64	1,348,315	3,448,737	2.56	1,335,900	3,404,667	2.55
대구	873,557	2,374,112	2.72	937,573	2,466,052	2.63	928,528	2,436,770	2.62
인천	920,471	2,560,167	2.78	1,066,297	2,890,451	2.71	1,045,417	2,822,601	2.70
광주	518,276	1,421,291	2.74	573,181	1,502,881	2.62	567,157	1,481,289	2.61
대전	535,947	1,439,526	2.69	588,395	1,538,394	2.61	582,504	1,519,314	2.61
울산	377,523	1,040,506	2.76	434,058	1,166,615	2.69	423,412	1,136,755	2.68
세종	45,095	108,496	2.41	76,419	204,088	2.67	75,219	199,617	2.65
경기	3,905,363	10,897,852	2.79	4,537,581	12,479,061	2.75	4,384,742	12,026,429	2.74
강원	560,024	1,390,197	2.48	611,578	1,518,040	2.48	606,117	1,499,734	2.47
충북	564,019	1,437,355	2.55	613,004	1,589,347	2.59	601,856	1,548,589	2.57
충남	756,381	1,897,648	2.51	816,247	2,107,802	2.58	796,185	2,036,720	2.56
전북	661,787	1,693,848	2.56	726,572	1,834,114	2.52	717,311	1,804,184	2.52
전남	641,290	1,579,118	2.46	730,743	1,799,044	2.46	720,612	1,764,433	2.45
경북	1,013,197	2,456,156	2.42	1,078,479	2,680,294	2.49	1,062,724	2,622,729	2.47
경남	1,160,110	3,009,319	2.59	1,282,617	3,334,524	2.60	1,258,487	3,244,163	2.58
제주	186,495	509,956	2.73	224,724	605,619	2.69	220,369	590,473	2.68

주 : 2010년 총조사 시점의 주소를 이용해 세종시의 가구 및 인구수를 산출함.

주택총조사 결과와 비교해 보면, 전체 인구의 도시 집중화, 남성의 생산연령인구, 읍면지역의 여성 고령화가 더 높아져 인구모집단의 성별, 연령대별 구조가 변화한 것으로 보인다.

<표 2-17>은 2010년 인구주택총조사와 2015년 인구주택총조사를 비교한 자료이다. 기본적으로 가구당 평균 가구원수가 감소하고, 젊은층의 청년 가구가 증가하여 약 206만 가구가 증가하였고, 인구는 약 402만 명이 증가하였다. 2015년 모집단은 2010년에 비해 대도시, 수도권 지역에서 많이 증가한 것으로 나타났다. 특히 세종은 2015년 모집단이 가장 많이 증가하였다.

이러한 결과로 볼 때, KLIPS의 추가표본은 현재 모집단의 구조 및 특성이 반영된 추출틀을 사용하여 표본을 추출하는 것이 필요하다.

나) 2015년 인구주택총조사의 추출틀

통계청에서는 인구주택총조사 이후 다양한 표본설계의 추출틀로 사용하도록 가구 기준으로 조사구를 제공한다. 통계청에서는 2015년 인구주택총조사의 20% 표본조사를 위해 2014년 12월 기준으로 가구주택기초조사를 실시하여 인구주택총조사를 위한 조사구를 구축하였다.

2014년 가구주택기초조사를 통해 구축한 조사구는 평균 30가구를 기준으로 획정하는데, 이는 과거 평균 60가구 기준의 조사구보다 조사구의 가구수를 1/2로 축소하였기에 조사구수는 약 2배 이상 증가되었다. 2015년 등록센서스 기반의 인구주택총조사 자료에 근거하여 2014년도에 구축한 평균 30가구 기준의 조사구를 보완하고 수정하여 2015년 인구주택총조사 기준의 조사구를 2017년 상반기(3~4월)에 공개할 예정이다. 그러나 2015년 등록센서스 기반의 인구주택총조사는 행정자료를 기반으로 조사구를 구축하기 때문에 과거 현장조사 기반의 조사구 작성에 비해 더 세심한 검토가 필요하므로 현재 통계청에서는 가구주택기초조사를 활용한 등록센서스 기반의 추출틀 수정 작업을 하고 있는 중이다.

<표 2-18>은 2015년 등록센서스 기반 인구주택총조사의 20% 표본조사를 위해 구축한 2014년 가구주택기초조사 기준의 조사구 자료이다. 2014년 가구주택기초조사를 통해 구축한 조사구는 평균 30가구를 단위로 획정하였고, 2015년 등록센서스 기반 인구주택총조사의 기본 추출틀로 활용할 것으로 파악되므로 이를 통해 2015년 인구주택총조사의 조사구를 예상할 수 있기에 살펴보기로 한다. <표 2-18>을 보면, 평균 30가구 기준으로 획정된 조사구는 모두 667,920개이다. 일반조사구(섬조사구 포함)는 354,418개, 아파트조사구는 313,502개이다. 일반적인 표본조사에서는 섬조사구를 제외하므로 실제 조사모집단 및 표본추출틀의 모집단 조사구수는 약간 감소한다. 경기도가 15만 개로 가장 많으며, 세종시는 2,402개로 가장 적다. <표 2-15>에서 보았듯이 대도시, 수도권의 가구수가 많으므로 조사구수도 대도시, 수도권이 많다.

〈표 2-18〉 가구주택기초조사의 조사구 규모(2014년 12월 기준)

	전체 조사구	아파트 조사구	기타 조사구	평균 가구수	읍면동수
전국	667,920	313,502	354,418	30.1	3,488
서울	128,703	53,234	75,469	30.1	423
부산	46,344	23,371	22,973	30.1	210
대구	32,056	16,290	15,766	30.1	139
인천	36,109	18,817	17,292	30.1	147
광주	19,090	12,076	7,014	30.1	95
대전	20,463	10,933	9,530	30.1	78
울산	14,687	7,945	6,742	30.1	56
세종	2,402	1,316	1,086	30.2	12
경기	150,312	80,480	69,832	30.1	552
강원	21,717	9,567	12,150	30.2	188
충북	21,964	9,828	12,136	30.2	153
충남	29,296	12,063	17,233	30.2	207
전북	26,137	11,683	14,454	30.2	241
전남	26,792	9,618	17,174	30.2	297
경북	39,261	14,114	25,147	30.2	332
경남	45,056	20,216	24,840	30.2	315
제주	7,531	1,951	5,580	30.2	43

주 : 기타조사구 : 일반조사구 + 섬조사구.

그러나 앞서 설명한 바와 같이 2015년에 실시된 등록센서스 기반의 인구주택총조사 자료에 근거하여 구축되는 평균 30가구 기준의 개별 조사구는 아직 수정 및 보완 작업을 하는 중이며, 2017년 상반기(3~4월)에 공개할 예정이라서 현시점에서 이용할 수 없는 실정이다.

3) KLIPS의 추가표본 추출을 위한 추출틀의 구축 방안

앞에서 설명한 바와 같이 KLIPS의 추가표본 추출을 위한 추출틀은 2015년 인구주택총조사의 조사구를 사용하는 것이 바람직하지만 2017년 상반기에 공개될 예정이라 현시점에 사용할 수 없어서 2010년 인구주택총조사의 조사구를 사용하기 위한 방안을 검토해 보았다.

첫째, 2010년 인구주택총조사의 조사구는 평균 60가구 기준으로 구축된 조사구가 공개된 상태이지만 2015년 현재 시점에서의 모집단과 비교

한 결과(표 2-17)를 보면, 현재 모집단의 구조와 특성이 매우 다른 것으로 나타나 시의성, 충분성이 부족하고 현재 모집단의 변화를 충분히 반영하지 못하는 것으로 파악된다. 2010년 인구주택총조사의 조사구를 추출틀로 사용하기 위해서는 가구 관점에서는 가구당 인구 감소, 1인 가구 증가(도시/청년층) 및 3인 이상 가구 감소를, 인구 관점에서는 고령층 증가, 청소년층 구성비 감소, 인구 출생률 저하, 수도권 및 도시 집중화 증가, 거주형태 면에서는 아파트 및 공동주택(다세대 및 주상복합/오피스텔 등) 거주 비율 증가 등 모집단의 변화된 특성과 구조가 충분히 반영되도록 수정하는 것이 필요하다. 현실적으로 가구 및 인구는 변화된 모집단 구조 및 특성을 현재 모집단을 통해 파악할 수 있지만 표본 추출을 위한 2010년 조사구에 2010년 이후 변화된 특성이 반영되도록 수정하거나 보완을 할 수는 없는 상황이다.

둘째, 2010년 총조사의 추출틀을 활용하기 위해서는 최신 모집단 정보를 벤치마크 자료로 활용해 추출틀을 보완하는 방안이 가능하다. 총조사 이후 모집단에서 거주형태의 변화를 반영하기 위한 벤치마크 자료로 신축아파트의 단지 및 세대별 자료를 2010년 추출틀(조사구)에 추가하는 방안을 검토해 보았다. 신축아파트 추출틀은 총조사 이후 신축아파트(준공기준) 단지 명부를 활용하여 기존 조사구와 유사한 가구규모(평균 60가구)를 유지하도록 가상의 조사구를 생성하여 활용하도록 신축아파트 조사구를 만들게 된다. 즉 추가표본 추출을 위한 추출틀로 2010년 인구주택총조사의 조사구와 신축아파트 조사구를 결합하여 추출틀로 사용하는 방안이다. 그러나 이 추출틀은 신축아파트 단지 정보를 수집하고 가상의 조사구를 구축하는 데 많은 시간과 비용이 소요되며, 또한 단독, 연립 등 일반조사구에 대한 조사구의 조정을 못하고, 2010년 조사구는 너무 오래전 추출틀이기 때문에 현시점에서 KLIPS의 추가표본 추출을 위한 추출틀로는 적절하지 않다.

결론적으로 추가표본을 위한 추출틀의 검토 결과를 정리하면, KLIPS의 추가표본 추출을 위한 추출틀은 2015년 인구주택총조사의 조사구를 사용하는 것이 가장 바람직하지만 2015년 인구주택총조사의 개별 조사구는 2017년 상반기에 공개될 예정이므로, 현시점에서 KLIPS의 추가표본

을 위한 추출틀로 2015년 인구주택총조사의 조사구를 추출틀로 활용할 수 없는 실정이다.

따라서 KLIPS의 추가표본은 ① 추가표본의 추출을 위한 추출틀로 2017년 상반기에 공개되는 2015년 인구주택총조사의 최신 추출틀을 사용하기 위해 표본설계 시점을 2017년 상반기로 연기하거나 ② 현재 시점에서 표본 추가에 필요한 표본설계를 수행해야 한다면, 2010년 인구주택조사구와 벤치마크한 신축아파트 단지 명부를 결합해 수정한 추출틀을 사용하는 방안 중 하나를 선택해야 한다. 그러나 추가표본은 현재 모집단 구조를 반영하고, 현행 표본의 마모 보완 및 변화된 표본 구조의 개선 등을 고려한다면 2017년 상반기로 표본설계 시점을 늦추어 2015년 인구주택총조사의 조사구를 사용하는 것이 바람직하다고 제언한다.

나. KLIPS 결합표본의 가중치

일반적으로 표본조사의 기본적인 표본설계는 모집단 정의 및 표본틀 설정, 층화, 표본배분, 표본추출, 가중치 산출 및 추정식 표현 등의 과정을 요구한다. 본 연구에서 제안하는 KLIPS의 추가표본을 위한 표본설계도 향후 결합표본의 관리, 가중치 및 추정치 산출 과정을 고려하면 KLIPS의 표본설계 틀을 유지하는 것이 중요하므로 현행 KLIPS의 표본설계와 동일한 표본설계의 틀을 유지하도록 한다. 추가표본 추출을 위한 표본설계에서 모집단과 표본틀은 현재 모집단과 최신 표본틀로 설정하고, 표본틀을 이용하여 층화, 층별 표본배분 및 표본추출의 과정은 현행 표본의 특성을 반영하여 KLIPS의 현행 표본설계와 동일하게 적용하면 된다. 여기서 표본 규모는 표본 확대 필요성을 고려하여 현행 패널과는 별도로 표본 규모를 결정하는 것이 바람직하며, 이때 현행 표본에서 표본 규모가 부족한 특정 계층(지역, 여성, 청년, 비정규직 등)은 과대표집으로 표본 규모를 결정하는 것이 필요하다. 실제로 표본틀을 설정한 후 수행하는 모집단 층화, 층별 표본배분 및 표본추출 과정은 실제 표본추출 시점에서 연구하면 되기 때문에 생략하고, 결합표본의 가중치 산출 방안과 그 과정을 살펴보고자 한다.

특히 호주의 HILDA 조사는 설계 시점에 비해 모집단이 증가하고, 모집단 구조가 변화하는 동시에 장기간 추적조사를 수행하기 때문에 조사거부, 추적실패 등의 표본 마모로 인한 표본구조가 모집단 구조와 차이가 발생하므로, 모집단 대표성을 확보하기 위해 원표본 기반의 현행 표본에 새로운 모집단에서 추출된 표본을 추가한 결합표본을 구축하는 방식으로 표본을 확대하고 있어 KLIPS의 표본 확대 방안과 유사하다.

그래서 본항에서는 KLIPS의 결합표본에 대한 가중치 산출을 위해 HILDA 조사의 결합표본에 대한 가중치 산출 방법과 그 과정을 소개한다.

1) HILDA의 횡단 및 종단 가중치 산출 과정

HILDA의 횡단 및 종단 가중치는 표본가구나 개인을 대상으로 표본 선택에 대한 설계 가중치, 응답 조정 그룹 혹은 모형으로부터 추정한 응답확률로 보정한 무응답보정 가중치, 가중치 추정 결과의 신뢰성 혹은 일관성을 높이기 위한 벤치마크 자료를 통한 캘리브레이션 사후층화 조정 가중치를 반영하여 최종 가중치를 결정한다.

종단가중치는 종단 연구를 위해 최소 1명 이상 응답한 응답가구 내에서 동거하고 있는 가구원(the enumerated person)과 응답자 모두에게 개인(the responding person) 종단가중치를 부여하지만, 가구 특성이 변화되고, 종단 가구를 이용하는 특정한 많은 연구 때문에 가구 종단가중치는 부여하지 않는다.

횡단가중치는 가구, 응답한 모든 개인과 응답가구에 동거하는 가구원에게 부여한다. 표본가구에 새로 유입해 TSM으로 추가되는 개인(동거인)은 횡단 추정에 포함하기 위해 횡단가중치만 부여한다.

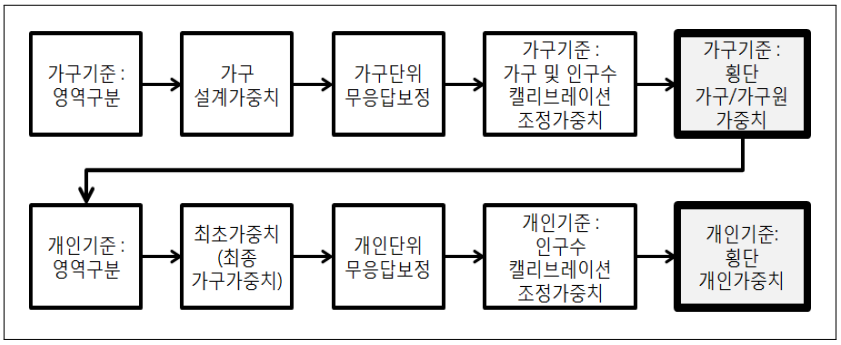
가) 횡단가중치

(1) 1차 조사의 횡단가중치

HILDA에서는 조사에 참여한 모든 표본을 대상으로 횡단가중치를 부여하는데, 1차 조사의 횡단가중치 산출 과정은 [그림 2-7]과 같다.

1) HILDA의 가구 및 가구원 횡단가중치는 적격 표본 가구를 대상으로 가구 단위에서 가구의 추출과정을 반영한 설계 가중치, 가구 단위의 실사

〔그림 2-7〕 HILDA의 1차조사 횡단가중치 산출 과정



결과에서 무응답 가구를 보정한 무응답보정 가중치를 산출한 후 사후층화 조정 가중치로 가구 및 인구수(개인 모집단)에 대한 벤치마크 자료를 이용하여 구한 캘리브레이션 조정 가중치를 곱해 횡단 가구 및 가구원 가중치를 부여한다. 이때 가구원 횡단가중치는 원가구원과 추가된 원가구원에게만 부여한다.

가구 및 가구원 횡단가중치를 위한 무응답 가구에 대한 보정은 주택 및 주거 특성(주택유형, 주택의 외형적인 조건, 주택의 보안 특성, 지역의 고층빌딩 비율)과 지역특성(지리적인 위치, 인구밀도, 제2외국어 사용비율, 비경제활동인구 비율, 실업인구비율, SEIFA 지수 등)을 이용한 로지스틱회귀모형으로부터 가구의 응답률을 추정해 무응답 가구에 대한 무응답보정 가중치를 산출한다.

HILDA의 벤치마크 자료로는 대상 모집단에 대한 총조사자료나 추계자료(호주 통계청의 추계 가구자료)를 이용하는데, 벤치마크 자료에 대해 조정한 캘리브레이션 조정 가중치를 결합해 가구 기준의 횡단 가구 및 가구원 가중치를 부여한다. HILDA에서 사용하는 벤치마크 자료는 <표 2-19>와 같다.

가구 및 가구원 횡단가중치 조정을 위한 HILDA의 벤치마크 자료를 보면, 개인기준 가중치 산출을 위한 주(state), 성별 및 연령별 벤치마크 자료로는 출생, 사망, 국내 및 국가 간 인구이동을 반영한 호주 통계청(ABS)의 추정인구를 이용한다. 추가해 가구 구성(자녀수)을 고려한 개인

〈표 2-19〉 HILDA의 벤치마크 기준

<table>
<tr><th></th><th>가구 가중치</th><th>가구원 가중치</th><th>개인 가중치</th></tr>
<tr><td rowspan="2">횡단가중치</td><td rowspan="2">성인수 × 아동수
주 × 규모</td><td>성별 × 연령대별
주 × 규모
경제활동 상태
결혼상태</td><td rowspan="2">성별 × 연령대별
주 × 규모
경제활동 상태
결혼상태
가구 구조
(성인 및 아동의 수)</td></tr>
<tr><td colspan="2">가구 및 동거인 가중치를 공동으로 결정</td></tr>
<tr><td>종단가중치</td><td>미제공</td><td>성별 × 연령대별
주 × 규모
경제활동 상태
결혼상태
가구 구조
(성인 및 아동의 수)</td><td>성별 × 연령대별
주 × 규모
경제활동 상태
결혼상태
가구 구조
(성인 및 아동의 수)</td></tr>
</table>

벤치마크는 가구 벤치마크로부터 계산하고, 경활(노동력) 및 결혼 상태의 벤치마크 자료로는 ABS의 노동력조사 자료를 이용한다. 벤치마크 자료는 '0' 가중치를 갖는 경우를 제외하기 위해 도서지역을 제외하거나 주거용이 아닌 주택에 거주하는 거주자를 제외하고 대상 모집단과 동일한 구조로 벤치마크 자료를 만든다.

2) HILDA의 개인 가중치는 가구 내 동거인의 적격 표본 여부에 관계없이 조사에 참여한 모든 개인을 대상으로 최종 횡단 가구가중치를 이용한 설계 가중치(가구의 설계 가중치), 무응답보정 및 캘리브레이션 조정 가중치를 결합해 개인 횡단가중치를 부여한다.

개인 가중치의 산출 과정은 가구기준의 가중치 산출 과정과 동일하지만 ① 초기 가중치는 최종 가구 가중치로 부여하고, ② 무응답자에 대한 보정은 2인 이상 개인이 있는 적격가구의 개인을 대상으로 무응답을 보정하는데, 가구의 주택 및 주거 특성과 지리적인 위치, 경활 상태, 성별, 연령대, 성인의 수, 아동의 수, 결혼 상태, 영어 사용 능력, 주거 유형 등을 이용한 로지스틱회귀모형으로부터 개인의 응답률을 추정해 무응답자에 대한 무응답보정 가중치를 계산하고, ③ 사후층화 조정을 위한 벤치마크 자료는 가구 내 동거인에 대한 벤치마크 자료와 기본적으로 동일하며, 전

체 응답자를 위한 캘리브레이션 조정 가중치의 계산은 가구 내 동거인에 대한 벤치마크와 동일한 방법으로 계산한다. 이 과정을 통해 기본 가중치, 무응답조정 및 켈리브레이션 조정 가중치를 결합해 응답한 전체 응답자에게 개인 횡단가중치를 부여한다.

(2) 2차 이후 횡단가중치

2차 조사 이후 10차 조사(2010년)까지 횡단 분석을 위해서 종단 표본만을 사용하는 것이 적절하지 않으므로 각 조사시점에서는 1차 조사 이후 변화된 모집단 특성과 모집단 커버리지를 증가시키고, 또한 변화된 응답 표본의 구조를 횡단 모집단으로 반영해 분석하는 것이 필요하므로 모든 응답자를 대상으로 모든 조사시점마다 횡단가중치를 부여한다.

HILDA의 가구 및 가구원, 개인에 대한 횡단가중치는 원표본 가구원(PSM)과 출생/분가 등 원표본으로 유입된 추가된 원표본(+PSM)과 함께 생활하는 일시적 조사대상이 되는 임시표본(TSM)을 포함하는 모든 개인을 대상으로 횡단가중치를 부여한다.

2차 이후 횡단가중치는 가구는 t시점에서 최소 1명 이상이 응답한 가구를 기준으로 부여하고, 가구원은 응답 가구에 속한 모든 응답자(TSM 포함)에게 부여한다. 응답자 중 사망, 해외이주 및 TSM에서의 이탈자는 무응답으로 제외하고, 응답한 모든 응답자에게 부여한다.

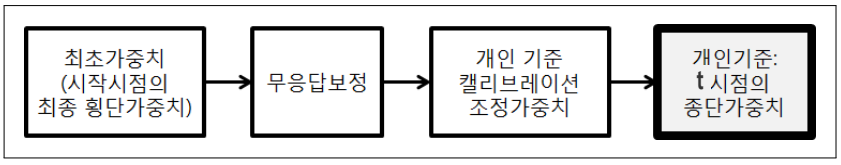
t시점에서의 가구 및 가구원 횡단가중치의 경우, t시점 조사에 참여한 응답자가 속한 가구가 t차 조사에 포함되는 다양한 경로를 반영한다. t시점에서의 가구 포함확률은 1차 조사의 가구 추출확률에 근거하여 t시점에 포함된 다양한 경로를 반영하여 추정한 포함확률을 계산하여 기본 가중치를 산출하고, 새로 유입된 응답자의 영향을 반영하기 위해 1차 조사의 가구 특성과 가구원 특성, 분가여부, 1차 이후 변화된 특성을 고려하여 가구에서의 무응답 보정을 위한 가구 유지 확률을 추정한 잠정 가중치를 산출해 무응답 보정을 하고, 잠정 가중치에 대한 캘리브레이션 조정을 통해 각 시점마다 가구 및 가구원의 최종 횡단가중치를 부여한다. t시점에서의 개인 가중치는 각 시점마다 최종 가구 횡단가중치를 기본가중치로 사용하여 HILDA의 개인 횡단가중치 산출 과정을 동일하게 적용해 부여한다.

나) 종단가중치

패널조사의 경우 기본적으로 많은 연구자는 1차조사 응답자 중 t시점까지 연속적으로 계속 참여한 응답자들의 종단 분석에 관심을 두므로 종단가중치를 부여해 분석한다.

HILDA의 개인 종단가중치를 보면, 1차 연도의 종단가중치는 개인 최종 횡단가중치와 동일하게 부여하지만 2차 이후 종단가중치는 매 조사시점마다 응답자 중 응답, 해외이주 및 사망으로 분류된 대상자를 계속 응답이 가능한 대상자로 정의하여, 원표본 및 임시표본에 관계없이 1차 조사에 참여한 모든 응답자 중 분석 시점까지 연속해 계속 참여한 응답자에게 종단가중치를 부여한다. 이 과정은 가구원 종단가중치의 산출 과정과 동일하며, 또한 임의의 a조사시점부터 b조사시점까지 종단 분석을 위해 종단가중치를 부여하거나 독립적인 임의의 두 시점 이상 비교를 위해 종단가중치를 부여할 때도 동일하게 적용한다.

〔그림 2-8〕 HILDA의 종단가중치 산출 과정



HILDA의 종단가중치를 부여하는 과정을 간략하게 살펴보면, 기본 가중치로는 종단 분석의 시작시점(원표본은 1차 기준)의 최종 횡단가중치를 부여하고, 시작시점부터 분석시점까지의 종단 표본 마모를 반영하기 위해 응답률에 영향을 주는 개인의 인구사회경제적 특성, 주택 및 주거특성, 가구 특성에 대한 시작시점의 자료와 분석시점까지의 변화된 특성뿐만 아니라 면접 상황을 기록한 파라데이터 특성을 반영한 로지스틱회귀모형을 통해 응답률을 추정해 무응답을 보정한 후, HILDA의 벤치마크 자료를 기준으로 캘리브레이션 조정을 통해 최종 종단가중치를 산출해 부여한다.

하지만 HILDA는 가구원 종단가중치와 개인 종단가중치를 부여하지만 가구 종단가중치는 제공하지 않는다.

다) 결합표본에서의 가중치

2001년부터 시작한 HILDA는 매 시점의 횡단 분석 결과에서 1차 조사 이후 해외에서 출생했거나 해외에서 국내로 이주한 15세 이상 인구 비율에 대한 횡단 추정치가 모집단 자료로 사용하는 노동력조사 추정치와의 차이가 점점 증가함으로써 이주자의 커버리지가 감소하여 횡단 추정의 편향이 매우 큰 변동을 보이는 것으로 파악하고 있다.

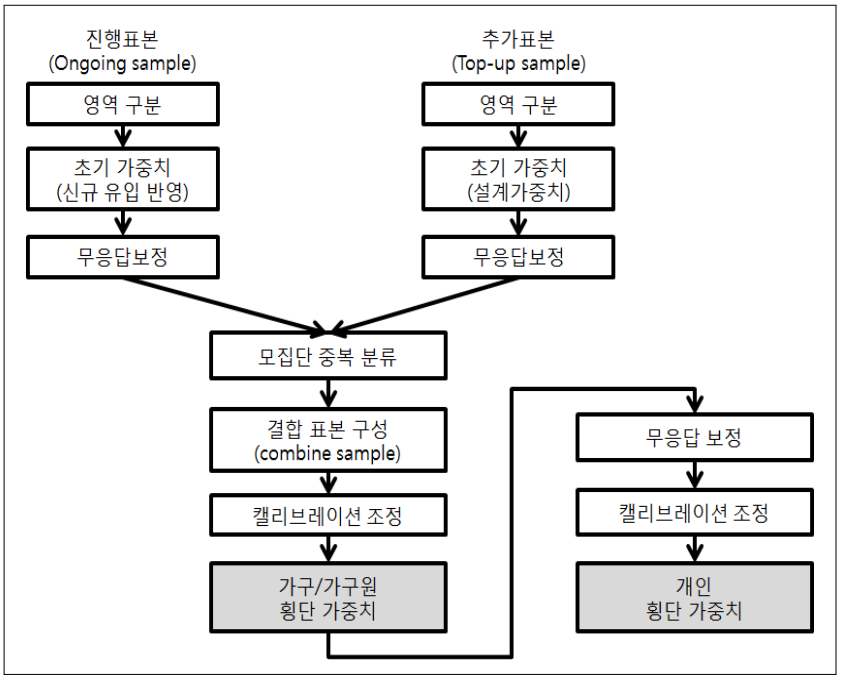
HILDA는 2011년에 1차 조사 이후 변화된 모집단 특성을 반영해 모집단 커버리지를 증가시키고, 또한 무응답 및 표본 마모로 인해 모집단 특성 구조와 달라진 표본의 특성 구조를 보정하기 위해 표본을 추가(Top-up sample)하는 방식으로 표본을 보완한다. 2011년 HILDA 표본은 2001년 최초 추출된 원표본 기반의 현재 표본에 2011년(11차 조사) 모집단에서 추출한 추가표본을 결합해 결합표본(혹은 통합표본, the integrated sample)으로 확대한다.

(1) 결합표본의 가중치 산출 과정

11차 조사의 HILDA 표본 가중치는 [그림 2-9]에서 보듯이 2001년 원표본 기반 표본과 2011년 새로 추가된 추가표본이 결합되어 있으므로 결합표본의 가중치 산출 과정이 복잡하다.

2011년 결합표본에 대한 HILDA의 횡단가중치 산출 과정을 간단하게 살펴보면, ① 결합표본에 대한 가중치 산출을 위해서 1차 조사로 설계된 진행표본(the ongoing sample)과 2011년의 추가표본(the top-up sample)을 분리해 현재의 진행표본에 대한 횡단가중치는 임시표본(TSM)을 포함한 모든 가구원에 대해 현행 가중치 산출 과정으로 산출하고, 2011년에 추가한 추가표본의 횡단가중치도 기본적으로 현행 표본의 가중치 부여 과정을 적용하여 설계 가중치 및 무응답보정 가중치를 산출하고, ② 캘리브레이션 조정 가중치를 산출하기 전에 두 표본의 구조적인 특성을 반영하여 결합표본을 구성해 결합표본에 대한 횡단가중치를 산출하고, ③ ②의 과정으로 계산한 결합표본의 가구 및 가구원에 대한 가중치를 현재 모집단 기준으로 캘리브레이션 조정을 통해 가구 및 가구원의 최종 횡단가중치를 산출하고, ④ 결합표본의 최종 가구 횡단가중치를 기본 가중치로

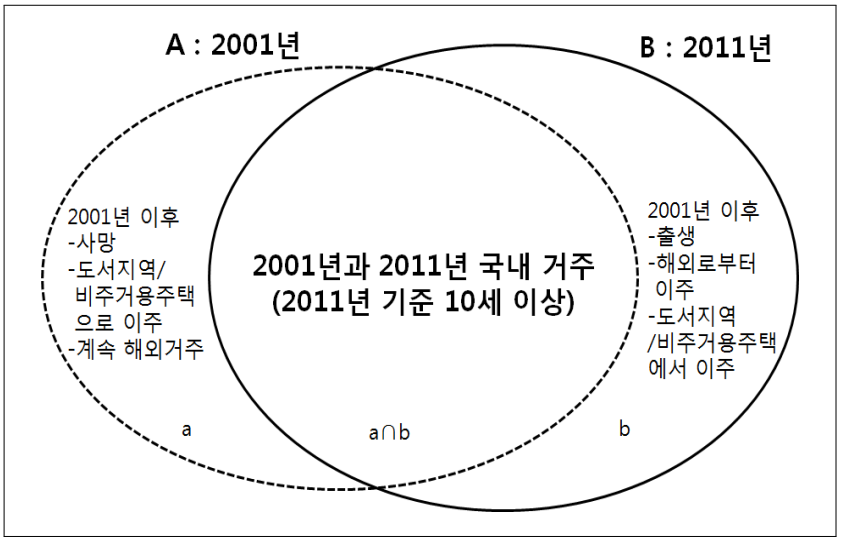
〔그림 2-9〕 결합표본의 가중치 산출 과정



이용하여 11차 조사에 응답한 모든 개인에게 부여하는 개인 최종 횡단가중치를 산출한다. 결합표본의 개인 횡단가중치 산출 방법은 현행 횡단가중치 산출 방법과 동일하게 적용한다.

11차 조사에서 현행 표본과 추가표본을 결합한 결합표본을 구성할 때, 두 설계 시점의 조사 모집단의 차이를 반영하는 것이 필요하다. [그림 2-10]에서 보듯이 2001년과 2011년의 조사모집단은 차이가 있다. 결합표본의 2011년 조사모집단을 1차 조사(2001년)의 조사모집단과 비교해 보면, ①2001년과 2011년의 두 조사모집단에 모두 포함된 조사대상(A & B), ②2001년 이후 사망했거나, 해외 이주 후 귀국하지 않거나 또는 1차 조사 이후 도서지역 혹은 비주거용 주택으로 이주하여 2011년 조사모집단에서 제외된 인구, ③2001년 이후 출생한 10세 이하의 아동, 해외로부터 이주하거나 1차 조사 이후 도서지역 혹은 비주거용 주택에서 조사대상 지역으로 이주하여 2011년 조사모집단에 포함된 인구 등으로 분류가

〔그림 2-10〕 추가표본의 모집단 구조와 유형



가능하다.

2011년에 추출한 추가표본은 개인보다는 개인의 가구를 기준으로 2011년 조사 모집단에서 어느 분류에 속하는지 구분한다. 2001년 원표본 기준의 현행 표본에서 제외된 도서지역이나 비주거용 주택의 가구가 2011년 추가표본으로 추출되는 경우는 추가표본의 가중치를 부여하지만, 가구원의 변동이 없거나 일부만 있는 가구는 결합표본 가중치를 부여한다. 그리고 2001년 원표본 가구원 중 사망 혹은 해외 이주 개인은 이미 횡단가중치 산출 시 제외되었으므로 반영하지 않는다.

〈표 2-20〉 11차(2011년)조사에서 결합표본의 표본유형별 가구수

		결합 사유	원표본	추가표본	비고
현행표본	a	비거주용 주택으로 이주	85	-	횡단가중치 제외
		도서지역으로 이주	28	-	횡단가중치 제외
현행표본 추가표본	a∩b	가구원 변동이 없는 가구	7,104	1,884	결합 가중치
		가구원이 일부 유입한 가구	149	90	결합 가중치
추가표본	b	전체 가구원이 유입한 가구	24	179	추가표본 가중치
총 가구수			7,390	2,153	

(2) 결합표본의 가중치 산출 방법

이 과정은 결합표본 가중치에 대해 캘리브레이션 조정 가중치를 반영하기 전 과정으로, 두 표본을 결합한 결합표본의 가중치 통합 방법 및 산출 과정을 소개한다.

HILDA는 2011년에 표본을 추가함으로써 현행 표본과 2011년에 추가한 추가표본을 결합한 결합표본으로 구성되어 있다. HILDA는 2001년 시작된 이래 변화된 모집단 구조와 마모된 표본에 대해 현재 모집단의 대표성을 확보하기 위해 11차(2011년도) 조사에서 추가표본을 추출하여 원표본 기준의 현행 표본과 합하여 결합표본(the pooling sample)으로 확대하고, 두 표본 추정결과를 합하여 결합 추정치(the combined estimator)를 생산한다.

(가) 결합 추정을 고려한 결합표본의 가중치

일반적으로 두 표본을 결합한 결합표본에서의 추정치는 두 표본으로부터 구한 추정치의 가중평균으로 추정한다. 두 표본 A, B로부터 관심변수 Y의 총합(total)에 대한 결합 추정치(the combined estimator)는 다음과 같다(O'Muircheartaigh & Pedlow, 2002).

$$\hat{Y}_{combined} = \theta \hat{Y}_A + (1-\theta)\hat{Y}_B$$

여기서 $0 \leq \theta \leq 1$, $\hat{Y}_A$, $\hat{Y}_B$는 표본 A, B의 추정치이다.

만약 두 표본이 독립이라면, 추정치 $\hat{Y}_{combined}$의 분산을 최소화하는 최적 θ는 다음과 같다. θ는 결합표본에서의 각 표본에 대한 포함률이며, 표본 배분률(a panel fraction factor)과 같은 의미이다.

$$\theta = \frac{\dfrac{n_A}{\mathrm{deff_A}}}{\dfrac{n_A}{\mathrm{deff_A}} + \dfrac{n_B}{\mathrm{deff_B}}}$$

여기서 n_A, $\mathrm{deff_A}$는 표본 A의 표본크기와 추정치의 설계효과이고, n_B, $\mathrm{deff_B}$는 표본 B의 표본크기와 추정치의 설계효과이다.

그리고 효율적인 추정 관점에서 두 표본을 결합한 결합표본(the combined sample)에서의 가중치는 다음과 같다.

$$w_{i,combined} = \begin{cases} \theta w_{iA} & , \text{if } i \in S_A \\ (1-\theta) w_{iB} & , \text{if } i \in S_B \end{cases}$$

여기서 w_{iA}, w_{iB}는 표본 A와 B에서의 개별 표본의 최종 가중치이다.

HILDA의 경우, 11차(2011년) 조사에 추가표본을 추출하여 원표본과 결합한 결합표본의 최종 횡단가중치를 표본그룹과 표본유형에 따라 각 표본의 가중치를 이용해 계산하면 다음과 같다.

$$w_{i,combined} = \begin{cases} 0 & , \text{if } i \in S_{A,a} \\ \theta w_{iA} & , \text{if } i \in S_{A,ab} \\ (1-\theta) w_{iB} & , \text{if } i \in S_{B,ab} \\ 0 & , \text{if } i \in S_{A,b} \\ w_{iB} & , \text{if } i \in S_{B,b} \end{cases}$$

여기서 w_{iA}는 현행 표본(A)에서 무응답자와 임시표본(TSM)에 대해 보정한 현행 표본의 가중치이고, w_{iB}는 2011년 추가표본(B)에서 무응답자 보정 후 추가표본의 가중치이다.

(나) 표본의 결합 특성을 고려한 결합표본의 가중치

그러나 결합표본에서의 추정치는 θ의 영향을 받게 되므로 결합표본에서는 θ의 최적 값을 결정하는 것이 매우 중요하다. HILDA는 모집단의 중복률을 고려하여 개별적인 두 표본의 결합 비율 θ를 결정하는데, 결합 추정치의 분산을 감소시키면서 편향을 최소화하도록 결정한다. 이를 위해 HILDA에서는 두 표본을 결합하고, 각 표본에서 표본으로 추출되고 응답하는 추출과 응답을 동시에 고려하는 추출과 응답의 결합확률(the probability of selection and response)을 추정하는 것이 추정치에 대한 최선의 개선방법이라고 판단하고 있다.

2011년 결합 모집단에서 가구원 변동이 없는 가구와 가구원의 일부 변동이 있는 가구를 포함한 결합 대상의 가구(유형a∩b)에 대한 병합 가중치(the pooled weight)는 현행 표본과 추가표본에서 표본 추출과 응답의 결합확률을 알고 있으므로 다음과 같이 계산한다.

$$w_{11h,pooled} = \frac{1}{p_{iA} + p_{iB}}$$

여기서 p_{iA}는 현행 표본(A)이 11차(2011년) 조사에서 표본으로 추출되고 응답하는 결합확률로 신규 유입과 무응답 보정 후 계산된 잠정 가중치($w_{11h,\mathrm{interim}}$)의 역수로 계산하고, p_{iB}는 2011년의 추가표본(B)으로 추출되어 응답하는 결합확률로 무응답 보정 후 계산된 가중치($w_{11h,\mathrm{nonadj}}$)의 역수로 계산한다.

그러나 표본으로 추출되지 않은 대상이 조사에 참여해 표본으로 포함된 경우(유형 b)에는 표본 추출 확률을 알 수 없어서 표본추출과 응답의 결합확률을 계산할 수 없으므로 이들 표본에 대한 결합확률을 추정해 가중치를 부여해야 한다. HILDA의 경우, 원표본 아동(PSM)이 원가구를 이탈해 원표본 가구원이 아닌 성인들과 새로운 가구를 형성해 현행 표본으로 포함된다면 이 가구의 표본 추출 확률을 알 수 없으므로 추출과 응답의 결합확률은 알 수 없다. 이처럼 표본으로 추출하지 않은 가구가 개별조사의 표본으로 포함되면 해당 가구의 추출과 응답에 대한 결합확률을 추정하기 위해 새로 포함된 가구원이 있는 가구의 가중치를 조정하는 방법으로 널리 사용하는 방법처럼 추출되지 않은 표본이 현행 표본(A)으로 포함되면 2011년 추가표본(B)에 대한 회귀모형으로부터 추출과 응답의 결합확률을 추정하고, 반대로 추출되지 않은 표본이 2011년 추가표본(B)으로 포함되면 현행 표본(A)의 회귀모형으로부터 추출과 응답의 결합확률을 추정한다. 즉 현행 표본(A) 중 2011년 조사에서 전체 가구원이 새로 유입된 가구(유형 b)는 현행 표본으로 추출된 가구가 아니므로 현행 표본에서 가구의 표본 추출과 응답의 결합확률을 알 수 없기 때문에 2011년 추가표본(B)으로부터 구한 추출과 응답의 결합확률에 대한 회귀모형에 현행 표본으로 신규 유입된 가구의 가구 특성과 가구주의 개인 특성을 대입하여 추정한 확률 $\hat{p}_{iB}$을 결합확률로 이용한다. 2011년 추가표본(B)에서도 표본으로 추출되지 않았지만 2011년의 추가표본에 포함된 가구(유형 b)는 추가표본으로 추출된 표본이 아니므로 표본 추출과 응답의 결합확률을 알 수 없기 때문에 현행 표본(A)에서 구한 추출과 응답에 대한 결

합확률의 회귀모형에 2011년의 추가표본으로 포함된 가구의 가구 특성과 가구주의 개인 특성을 대입하여 추정한 확률 $\hat{p}_{iA}$을 결합확률로 이용한다. 이를 이용하여 계산한 가중치는 다음과 같다.

$$w_{i,pooled} = \begin{cases} \dfrac{1}{p_{iA} + \hat{p}_{iB}} & , \text{if } i \in S_A \\ \dfrac{1}{\hat{p}_{iA} + p_{iB}} & , \text{if } i \in S_B \end{cases}$$

2011년에 추가표본을 추출하여 원표본과 결합한 결합표본에 대한 최종 횡단가중치를 표본그룹과 표본유형을 고려하여 병합가중치(the pooled weight)로 계산하면 다음과 같다.

$$w_{i,pooled} = \begin{cases} 0 & , \text{if } i \in S_{A,a} \\ \dfrac{1}{p_{iA} + \hat{p}_{iB}} & , \text{if } i \in S_{A,ab} \\ \dfrac{1}{\hat{p}_{iA} + p_{iB}} & , \text{if } i \in S_{B,ab} \\ 0 & , \text{if } i \in S_{A,b} \\ w_{iB} & , \text{if } i \in S_{B,b} \end{cases}$$

여기서 $p_{iA} = \dfrac{1}{w_{iA}}$, $p_{iB} = \dfrac{1}{w_{iB}}$는 현행 표본(A)과 2011년 추가표본(B)에 대한 표본 추출과 응답의 결합확률이고, $\hat{p}_{iA}$, $\hat{p}_{iB}$는 현행 표본(A)과 2011년 추가표본(B)에서 표본으로 추출되지 않은 표본의 추출과 응답에 대해 추정한 결합확률이다.

HILDA의 가중치 부여 과정 및 방법에 대한 자세한 사항은 “Longitudinal and Cross-sectional Weighting Methodology for the HILDA Survey (HILDA project technical paper series, No. 2/12, Dec. 2012, N. Watson)”을 참조하기 바란다.

2) KLIPS 결합표본의 가중치 산출 방안

KLIPS는 2015년도를 기준으로 보면, 98표본은 제18차, 09통합표본으로 7차조사로 ①98표본의 원표본가구 유지율은 2015년(18차) 조사에서 68.4%

를 유지하고, ②09통합표본의 원표본가구의 유지율은 86.2%이고, ③2015년 조사에서 소멸가구를 제외한 조사대상 가구는 모두 9,692가구(98표본 8,157가구, 추가표본 1,535가구)이며, 조사에 성공한 가구는 6,934가구이고, 가구원은 모두 13,373명이다.

본 연구에서는 KLIPS의 추가표본 방안으로 표본 이탈 및 표본 마모로 인해 변화된 표본구조의 대표성을 유지하고 변화된 모집단의 대표성을 확보하면서 지역 및 특정 계층의 통계 생산이 가능하도록, 현재 모집단에서 기존 설계틀을 유지하면서 표본을 추가하는 방안을 제안한다. 그 방안은 KLIPS의 98원표본과 09통합표본의 원패널에 현재시점(2016년 기준)의 모집단에서 추출된 표본을 추가한 결합표본(the combined sample)으로 확대하는 방안이다. 표본 관리는 09통합표본을 기준으로 변화된 모집단에서 표본을 추가적으로 추출해 현행 표본과 결합해 관리하는 방안이 적절하다.

본 연구에서는 KLIPS의 결합표본에 대한 가중치를 산출하는 방안으로 HILDA와 같이 표본가구나 개인을 대상으로 표본 선택에 대한 설계 가중치, 응답 조정 그룹 혹은 모형으로부터 추정한 응답 확률로 보정한 무응답보정 가중치, 가중치 추정 결과의 신뢰성 혹은 일관성을 높이기 위한 벤치마크 자료를 통한 캘리브레이션 사후층화 조정 가중치를 반영하여 최종 가중치를 산출하는 방안을 제안한다.

본 연구에서 소개한 HILDA의 가중치 산출 방법과 과정을 면밀히 검토하고, 2017년 상반기에 2015년 인구주택총조사의 조사구가 공개되어 추가표본을 추출한 후, 17 KLIPS 결합표본의 추정 및 분석을 위한 가중치 연구에서 추가적으로 검토하는 것이 필요하다.

제5절 추가 표본추출을 위한 2017년 시범조사 표본설계

본절에서는 현행 노동패널에 약 6,000가구를 추가한 2017통합패널(가칭) 구축을 위한 표본 설계 방안을 제시하고자 한다. 앞서 제4절에서 살

펴본 바와 같이 2017년통합패널 구축을 위한 추가표본은 2015년 인구주택총조사를 모집단으로 설정한다. 2017통합패널 구축의 큰 방향은 2015년 총조사를 모집단으로 설정하여 광역시도 및 읍면부, 주택유형별로 목표 표본크기를 배분하여 결정한 후, 2017통합패널에 포함하여 유지될 수 있는 현행 표본 규모를 제외한 부족한 표본 규모를 광역시도 및 읍면부, 주택유형별로 추가하기로 한다.

1. 2015 인구주택총조사의 특징

2015년 인구주택총조사는 2010년 이전의 총조사와 달리 다음과 같은 점에서 큰 특성을 가진다.

첫째, 2015년 인구주택총조사의 가장 큰 특징은 2010년까지 사용되던 현장조사 기반의 자료수집 체계가 행정자료 및 등록자료 기반의 등록센서스로 자료수집체계가 변경되었다는 점이다. 이는 가구 및 가구원, 주택 등의 전수 항목에 대해서 현장조사 기반의 실측자료 대신 주민등록명부, 주택대장 등 행정자료 기반의 등록자료로 자료를 수집하는 것이다.

둘째, 집계를 위한 기초자료 및 기준이 가구 단위에서 행정지역 단위로 변경되었다는 점이다. 현장조사 기반에서는 가구 단위를 기초자료로 하여 행정지역 단위로 집계하여 통계를 생산하였지만, 등록자료 기반에서는 소지역 추정법을 응용하여 전국 및 광역 단위 기준의 통계를 적성한 후 소지역의 행정지역 단위로 통계를 생산하는 것이다.

셋째, 표본추출틀 제공을 위한 조사구 구축에서도 등록센서스에서는 가구별 개별 자료가 존재하지 않아 가구 단위 기준의 조사구 구축이 불가능하므로 행정자료 기반으로 동읍면 단위에서 조사구를 구축하였다는 점이다. 즉 등록자료 기반의 조사구에서는 조사구 내 가구 및 가구원, 주택 등에 대한 집계 자료만을 제공하므로 가구 단위의 연계가 불가능하고 조사구 특성만을 파악할 수 있는 정보만으로 조사구를 구축한다는 것이다.

넷째, 2015년 조사구는 2014년 12월에 통계청에서 수행한 가구주택기초조사에 근거해 작성된 것이다. 가구주택기초조사는 기본적으로 모집단 변화를 파악하면서 행정자료 기반의 총조사를 위해 행정자료와 실측자료

와의 차이 검토를 통해 행정자료 보정 연구의 기초자료로 활용하고, 또한 60가구 기준의 2010년 조사구를 30가구 단위로 분할하기 위해 수행한 조사이다. 이 조사를 통해 구축한 조사구는 2015년 총조사의 표본조사를 위한 조사구 추출틀로 사용하였고, 2015년 총조사의 행정자료 및 등록자료로 수정 및 보완하여 2015년 조사구를 작성한 것이다.

다섯째, 2015년 기본 조사구의 가구 규모는 평균 30가구이고, 평균 60가구 기준의 병합조사구를 별도로 제공하고 있다. 2015년 조사구는 행정자료 기반으로 새롭게 조사구를 구축한 것이 아니라 2014년 가구주택기초조사의 조사구에 근거해 조사구의 소멸 및 개발 등 조사구의 변화와 조사구 내 가구의 변화를 반영해 평균 30가구 기준을 유지하도록 기본 조사구를 작성하고, 인접한 조사구를 결합하여 병합조사구를 구축한 것이다. 참고로, 일반가구 기준의 2014년 가구주택기초조사의 조사구는 보통조사구 354,418개, 아파트조사구 313,502개로 모두 667,920개였고, 2015년 총조사의 조사구는 보통조사구 353,253개, 아파트조사구 314,996개로 모두 668,249개이다. 단독주택, 연립 및 다세대 주택 등의 보통조사구는 감소하고, 아파트조사구는 증가하였다.

여섯째, 2015년 조사구의 가구수 규모 변동이 2010년에 비해 크다는 점이다. 평균 30가구 기준인 기본 조사구의 경우 20가구 미만이 12.4%이고, 50가구 이상이 1.4%이고, 평균 60가구 기준인 병합조사구의 경우에도 40가구 미만이 19.0%이고, 100가구 이상도 3.8%이다. 실제로 가구수가 1가구인 조사구도 존재하고, 1,000가구 이상 조사구도 존재한다.

2. 2017통합패널의 모집단과 표본틀

가. 모집단

한국노동패널의 목표 모집단은 국내에 거주하는 일반가구 및 가구원이고, 조사 모집단은 특수조사구를 제외하고 국내에 거주하는 일반가구 및 가구원이다.

<표 2-21>은 2015년 인구주택총조사 기준의 모집단 가구 규모를 나타

〈표 2-21〉 모집단 규모(광역시도별, 동/읍면부별 및 주택유형별 가구수)

		계	아파트	보통조사구					
				소계	단독	단독_이외			
						소계	거주_단독이외	비거주 주택	
								오피스텔	기타
전국	소계	19,111,030	9,196,769	9,914,261	6,738,920	3,175,341	2,461,355	320,193	393,793
	동부	15,487,901	8,148,304	7,339,597	4,590,772	2,748,825	2,179,085	312,261	257,479
	읍면부	3,623,129	1,048,465	2,574,664	2,148,148	426,516	282,270	7,932	136,314
서울	소계	3,784,490	1,586,400	2,198,090	1,187,440	1,010,650	816,425	115,000	79,225
	동부	3,784,490	1,586,400	2,198,090	1,187,440	1,010,650	816,425	115,000	79,225
	읍면부	0	0	0	0	0	0	0	0
부산	소계	1,335,900	691,595	644,305	413,601	230,704	175,331	37,072	18,301
	동부	1,281,894	655,380	626,514	402,117	224,397	170,651	36,512	17,234
	읍면부	54,006	36,215	17,791	11,484	6307	4,680	560	1,067
대구	소계	928,528	486,653	441,875	349,913	91,962	73,954	4,783	13,225
	동부	862,069	444,299	417,770	331,414	86,356	70,571	4,632	11,153
	읍면부	66,459	42,354	24,105	18,499	5606	3,383	151	2,072
인천	소계	1,045,417	555,329	490,088	204,266	285,822	242,237	26,890	16,695
	동부	1,011,279	553,017	458,262	179,033	279,229	237,363	26,880	14,986
	읍면부	34,138	2,312	31,826	25,233	6,593	4,874	10	1,709
광주	소계	567,157	355,697	211,460	176,387	35,073	20,689	4,770	9,614
	동부	567,157	355,697	211,460	176,387	35,073	20,689	4,770	9,614
	읍면부	0	0	0	0	0	0	0	0
대전	소계	582,504	318,020	264,484	204,479	60,005	47,295	5,473	7,237
	동부	582,504	318,020	264,484	204,479	60,005	47,295	5,473	7,237
	읍면부	0	0	0	0	0	0	0	0
울산	소계	423,412	237,732	185,680	141,375	44,305	35,444	2,918	5,943
	동부	345,357	195,621	149,736	112,955	36,781	29,884	2,732	4,165
	읍면부	78,055	42,111	35,944	28,420	7,524	5,560	186	1,778
세종	소계	75,219	46,921	28,298	23,366	4,932	2,524	855	1,553
	동부	34,919	33,873	1,046	59	987	12	823	152
	읍면부	40,300	13,048	27,252	23,307	3,945	2,512	32	1,401

〈표 2-21〉의 계속

		계	아파트	보통조사구					
				소계	단독	단독_이외			
						소계	거주_단독이외	비거주 주택	
								오피스텔	기타
경기	소계	4,384,742	2,410,360	1,974,382	1,139,402	834,980	652,408	86,538	96,034
	동부	3,640,067	2,115,840	1,524,227	845,061	679,166	536,289	83,763	59,114
	읍면부	744,675	294,520	450,155	294,341	155,814	116,119	2,775	36,920
강원	소계	606,117	261,233	344,884	290,093	54,791	36,973	999	16,819
	동부	361,671	206,653	155,018	129,496	25,522	19,034	883	5,605
	읍면부	244,446	54,580	189,866	160,597	29,269	17,939	116	11,214
충북	소계	601,856	274,002	327,854	273,310	54,544	39,008	2,236	13,300
	동부	352,646	203,199	149,447	122,278	27,169	21,751	1,375	4,043
	읍면부	249,210	70,803	178,407	151,032	27,375	17,257	861	9,257
충남	소계	796,185	338,990	457,195	370,954	86,241	58,835	6,793	20,613
	동부	349,263	207,812	141,451	100,696	40,755	28,243	6,083	6,429
	읍면부	446,922	131,178	315,744	270,258	45,486	30,592	710	14,184
전북	소계	717,311	327,120	390,191	341,020	49,171	32,906	1,270	14,995
	동부	493,420	295,307	198,113	164,091	34,022	25,819	1,201	7,002
	읍면부	223,891	31,813	192,078	176,929	15,149	7,087	69	7,993
전남	소계	720,612	261,576	459,036	402,877	56,159	32,513	4,564	19,082
	동부	297,136	182,784	114,352	89,314	25,038	16,558	3,469	5,011
	읍면부	423,476	78,792	344,684	313,563	31,121	15,955	1,095	14,071
경북	소계	1,062,724	401,582	661,142	551,089	110,053	79,492	2,851	27,710
	동부	553,228	291,888	261,340	198,210	63,130	50,815	2,500	9,815
	읍면부	509,496	109,694	399,802	352,879	46,923	28,677	351	17,895
경남	소계	1,258,487	587,673	670,814	557,999	112,815	73,032	14,061	25,722
	동부	809,718	448,947	360,771	283,120	77,651	52,550	13,078	12,023
	읍면부	448,769	138,726	310,043	274,879	35,164	20,482	983	13,699
제주	소계	220,369	55,886	164,483	111,349	53,134	42,289	3,120	7,725
	동부	161,083	53,567	107,516	64,622	42,894	35,136	3,087	4,671
	읍면부	59,286	2,319	56,967	46,727	10,240	7,153	33	3,054

〈표 2-22〉 광역시도별, 동/읍면부별 및 가구원수별 가구수

	가구원수						
	일반가구	1명	2명	3명	4명	5명 이상	평균
전국	19,111,030	5,203,440	4,993,818	4,100,979	3,588,931	1,223,862	2.5
동부	15,487,901	4,124,961	3,844,701	3,430,598	3,092,711	994,930	2.6
읍면부	3,623,129	1,078,479	1,149,117	670,381	496,220	228,932	2.3
서울	3,784,490	1,115,744	930,467	817,440	701,945	218,894	2.5
동부	3,784,490	1,115,744	930,467	817,440	701,945	218,894	
읍면부							
부산	1,335,900	361,749	366,048	296,742	238,031	73,330	2.5
동부	1,351,689	349,929	350,939	284,091	227,140	69,795	
읍면부	57,541	11,820	15,109	12,651	10,891	3,535	
대구	928,528	239,517	239,824	208,795	184,258	56,134	2.6
동부	913,647	225,255	222,156	193,386	169,694	51,578	
읍면부	71,015	14,262	17,668	15,409	14,564	4,556	
인천	1,045,417	243,678	265,079	245,135	220,538	70,987	2.6
동부	1,011,279	232,704	252,809	239,430	217,171	69,165	
읍면부	34,673	10,974	12,270	5,705	3,367	1,822	
광주	567,157	163,577	137,662	115,701	109,612	40,605	2.5
동부	567,157	163,577	137,662	115,701	109,612	40,605	
읍면부							
대전	582,504	169,391	140,603	122,088	112,055	38,367	2.5
동부	582,504	169,391	140,603	122,088	112,055	38,367	
읍면부							
울산	423,412	103,551	104,101	100,230	90,735	24,795	2.6
동부	345,357	84,587	83,685	82,672	74,514	19,899	
읍면부	81,564	18,964	20,416	17,558	16,221	4,896	
세종	75,219	21,889	18,010	15,219	14,560	5,541	2.5
동부	34,919	7,098	7,263	8,384	9,404	2,770	
읍면부	41,754	14,791	10,747	6,835	5,156	2,771	
경기	4,384,742	1,026,471	1,062,222	1,005,330	971,146	319,573	2.7
동부	3,640,067	837,033	851,255	851,104	838,376	262,299	2.7
읍면부	744,675	189,438	210,967	154,226	132,770	57,274	2.6

〈표 2-22〉의 계속

	가구원수						
	일반가구	1명	2명	3명	4명	5명 이상	평균
강원	606,117	189,379	178,470	116,384	87,182	34,702	2.4
동부	361,671	108,028	99,520	73,901	59,435	20,787	2.4
읍면부	244,446	81,351	78,950	42,483	27,747	13,915	2.2
충북	601,856	173,598	167,146	120,634	100,527	39,951	2.5
동부	352,646	95,248	88,338	76,088	68,435	24,537	2.6
읍면부	249,210	78,350	78,808	44,546	32,092	15,414	2.3
충남	796,185	234,513	226,658	153,585	128,382	53,047	2.4
동부	349,263	101,252	82,314	72,316	69,374	24,007	2.5
읍면부	446,922	133,261	144,344	81,269	59,008	29,040	2.4
전북	717,311	213,750	209,008	135,856	110,132	48,565	2.4
동부	493,420	141,408	129,126	99,525	88,370	34,991	2.5
읍면부	223,891	72,342	79,882	36,331	21,762	13,574	2.3
전남	720,612	218,864	227,975	131,454	96,940	45,379	2.4
동부	297,136	79,998	82,559	61,990	52,141	20,448	2.5
읍면부	423,476	138,866	145,416	69,464	44,799	24,931	2.2
경북	1,062,724	322,569	320,498	207395	156,364	55,898	2.4
동부	553,228	158,957	146,709	117,967	98,922	30,673	2.5
읍면부	509,496	163,612	173,789	89,428	57,442	25,225	2.2
경남	1,258,487	346,754	341,745	265,127	228,216	76,645	2.5
동부	809,718	212,497	199,390	181,669	165,727	50,435	2.6
읍면부	448,769	134,257	142,355	83,458	62,489	26,210	2.3
제주	220,369	58,446	58,302	43,864	38,308	21,449	2.6
동부	161,083	42,255	39,906	32,846	30,396	15,680	2.6
읍면부	59,286	16,191	18,396	11,018	7,912	5,769	2.5

낸 표이다. 모집단 가구수는 19,111,030가구이다. 광역시도별로는 경기도가 약 438.5만 가구로 가장 많으며, 세종이 7.5만 가구로 가장 적다. 동읍면부별로는 동부가 1,548.8만 가구이고, 읍면부가 362.3만 가구이다. 보통조사구가 991.4만 가구이고, 아파트조사구가 919.7만 가구이다.

<표 2-22>는 광역시도, 동읍면부별 및 가구원수별 모집단 자료이다. 평균 가구원수는 2.5명이다. 평균 가구원수는 경기도가 2.7명으로 가장 많다. 동읍면부별로는 동부가 2.6명, 읍면부가 2.3명이다.

가구원수는 1인 가구가 520.3만 가구로 가장 많다. 특히 동읍면부별로 보면 동부에서 1인 가구수가 412.5만 가구로 가장 많고, 읍면부에서는 2인 가구가 114.9만 가구로 가장 많다.

나. 표본틀

한국노동패널조사의 2017통합패널의 표본추출틀은 2015년 인구주택총조사의 일반가구 조사구이다. 일반가구 조사구는 특수조사구와 섬조사구를 제외한 단독주택, 연립주택 및 다세대 주택 등의 보통조사구와 아파트 조사구로 구성된다.

2017통합패널 구축을 위한 표본추출틀의 규모를 보면, 일반가구 중 특수조사구와 섬조사구 가구를 제외하고 표본추출틀에 포함된 가구수는 19,100,262가구이다. 가구 기준 모집단 포함률은 99.94%이다. 참고로, 표본틀에 포함된 일반 가구원은 48,435,095명이다.

1) 기본 조사구 현황

기본적으로 2015년 인구주택조사구는 평균 30가구 기준의 기본조사구로 구축되어 있다. 기본조사구에 대한 가구수 및 가구원수(인구수) 특성을 간단히 살펴보면 다음과 같다.

2015년 인구주택조사구의 기본조사구 기준 가구수 및 가구원수 특성을 요약한 <표 2-23>을 보면, 표본추출틀에 포함된 조사구는 모두 668,249개 조사구이다. 조사구 기준 평균 가구수는 28.6가구이고, 평균 가구원수는 72.5명이다. 가구수의 경우, 최소는 1개 가구, 최대 2,120가구를 갖는 조사구가 존재하며, 가구원수도 최소 1명, 최대 6,044명을 갖는 조사구가 있다.

<표 2-24>와 <표 2-25>는 2015년 인구주택조사구의 기본조사구 기준 가구수 및 가구원수 규모별 분포를 요약한 결과이다. <표 2-24>를 보면, 조사구의 가구수가 20~30가구 미만인 조사구가 44.0%로 가장 많으며,

30~40가구 미만인 조사구도 34.4%로 많다. 2015년 인구주택총조사 기본조사구의 가구수가 20~50가구 미만인 조사구가 86.1%이다. 그러나 20가구 미만인 조사구도 12.4%나 된다. <표 2-25>를 보면, 조사구의 가구원수가 60~80명 미만인 조사구가 28.2%로 가장 많으며, 40~60명 미만인 조사구가 22.2%, 80~100명 미만인 조사구가 22.0%의 순으로 많다. 2015년 인구주택총조사 기본조사구의 가구원수가 40~120명 미만인 조사구가 82.8%이다. 그러나 40명 미만인 조사구도 11.7%나 된다.

〈표 2-23〉 2015년 인구주택조사구의 조사구 기준 가구수 및 인구수 특성

	N	평균값	표준편차	최소값	최대값
가구수	668,249	28.58	11.98	1	2,120
인구수	668,249	72.48	35.11	1	6,044

〈표 2-24〉 2015년 인구주택조사구의 가구수 규모별 분포

	빈도	백분율	누적빈도	누적백분율
10가구 미만	13,998	2.09	13,998	2.09
10~20	68,758	10.29	82,756	12.38
20~30	294,207	44.03	376,963	56.41
30~40	229,907	34.40	606,870	90.81
40~50	51,785	7.75	658,655	98.56
50가구 이상	9,594	1.44	668,249	100.00

〈표 2-25〉 2015년 인구주택조사구의 인구수 규모별 분포

	빈도	백분율	누적빈도	누적백분율
20명 미만	19,646	2.94	19,646	2.94
20~40	58,273	8.72	77,919	11.66
40~60	147,990	22.15	225,909	33.81
60~80	188,311	28.18	414,220	61.99
80~100	146,964	21.99	561,184	83.98
100~120	69,972	10.47	631,156	94.45
120명 이상	37,093	5.55	668,249	100.00

〈표 2-26〉 2015년 인구주택조사구의 가구수 분포별 가구수 및 인구수 특성

	조사구수	항목	평균	표준편차	최소값	최대값
10가구 미만	13,998	가구수	5.06	2.76	1	9
		인구수	9.98	28.83	1	3,079
10~20	68,758	가구수	16.23	2.64	10	19
		인구수	36.44	17.44	10	2,367
20~30	294,207	가구수	24.83	2.78	20	29
		인구수	63.60	18.36	20	2,809
30~40	229,907	가구수	33.66	2.93	30	39
		인구수	87.49	24.10	30	5,468
40~50	51,785	가구수	42.03	2.44	40	49
		인구수	106.98	27.09	40	2,184
50가구 이상	9,594	가구수	72.24	60.22	50	2,120
		인구수	148.33	138.67	50	6,044

〈표 2-27〉 2015년 인구주택조사구의 인구수 분포별 가구수 및 인구수 특성

	조사구수	항목	평균	표준편차	최소값	최대값
20명 미만	19,646	가구수	7.51	4.45	1	19
		인구수	10.80	5.65	1	19
20~40	58,273	가구수	18.46	4.82	1	39
		인구수	31.54	5.54	20	39
40~60	147,990	가구수	24.07	5.08	1	59
		인구수	50.43	5.63	40	59
60~80	188,311	가구수	28.36	5.53	1	78
		인구수	69.54	5.73	60	79
80~100	146,964	가구수	32.17	5.75	3	93
		인구수	88.63	5.67	80	99
100~120	69,972	가구수	36.43	6.12	2	116
		인구수	108.06	5.66	100	119
120명 이상	37,093	가구수	45.75	33.74	1	2,120
		인구수	141.27	82.90	120	6,044

<표 2-26>과 <표 2-27>은 2015년 인구주택조사구의 기본조사구 기준 가구수 및 가구원수 특성을 요약한 결과이다. <표 2-26>의 가구수 규모별 결과에서 특이한 점은, 가구수 규모별 가구원수의 최대 가구원수를 보면 모든 규모에서 2,000명 이상의 조사구가 존재한다는 점이다. <표 2-27>의 가구원수 규모별 결과에서 특이한 점은, 가구수와 가구원수가 같은 1인 가구로 구성된 조사구가 존재한다는 점이다. 또한 120명 이상인 조사구에서는 평균 가구수가 45.8가구이고, 평균 가구원수가 141.3명이나 된다.

특히, 조사구별 가구수 및 가구원수를 보면, 최소 조사구는 1가구, 1명의 조사구가 존재하는 반면, 최대 조사구는 2,120가구이고, 가구원수가 6,044명이나 된다.

<표 2-28>은 2015년 인구주택조사구의 기본조사구 기준 광역시도별 가구수 및 가구원수 특성을 요약한 결과이다. 광역시도별 평균 가구수 및 가구원수를 보면, 조사구의 평균 가구수는 경기도가 30.1가구로 가장 많

〈표 2-28〉 2015년 인구주택조사구의 광역시도별 가구수 및 인구수 특성

	조사구수	항목	평균	표준편차	최소값	최대값
일반가구	668,249	가구수	28.58	11.98	1	2,120
		인구수	72.48	35.11	1	6,044
서울	133,091	가구수	29.11	14.57	1	2,120
		인구수	72.09	36.01	1	6,044
부산	47,062	가구수	28.30	11.83	1	1,065
		인구수	70.57	33.28	1	3,184
대구	31,662	가구수	29.23	9.49	1	423
		인구수	74.92	29.60	1	951
인천	36,073	가구수	28.91	10.28	1	479
		인구수	76.62	31.87	1	826
광주	19,117	가구수	29.54	12.87	1	908
		인구수	75.03	31.89	1	1,310
대전	20,285	가구수	28.57	13.70	1	1,235
		인구수	72.25	56.75	1	5,468

〈표 2-28〉의 계속

	조사구수	항목	평균	표준편차	최소값	최대값
울산	15,099	가구수	28.18	8.72	1	294
		인구수	73.55	28.88	1	1,060
세종	2,708	가구수	24.79	12.05	1	311
		인구수	63.45	31.81	1	481
경기	148,227	가구수	30.11	12.59	1	1,943
		인구수	80.54	37.18	1	2,678
강원	21,462	가구수	27.06	9.26	1	256
		인구수	64.24	26.79	1	645
충북	21,843	가구수	27.08	10.29	1	634
		인구수	67.23	30.64	1	1,470
충남	28,279	가구수	27.69	12.50	1	880
		인구수	68.08	31.72	1	1,247
전북	25,604	가구수	27.34	9.45	1	497
		인구수	66.49	27.10	1	1,172
전남	24,600	가구수	27.30	8.78	1	341
		인구수	65.17	25.40	1	683
경북	39,315	가구수	26.44	10.14	1	672
		인구수	63.16	37.47	1	3,079
경남	45,987	가구수	27.05	9.13	1	439
		인구수	68.17	32.64	1	2,367
제주	7,835	가구수	27.60	9.94	1	262
		인구수	72.34	27.89	1	615

으며, 세종이 24.8가구로 가장 적다. 이를 제외한 나머지 지역은 모두 27~29가구로 나타나 광역시도 간 조사구의 평균 가구수 차이는 크지 않다. 조사구의 가구원수는 경기도가 80.5명으로 가장 많으며, 경북이 63.2명으로 가장 적고, 세종시와 강원도도 65명 이하로 나타나 다른 광역시도에 비해 조사구 내 가구원수가 적은 지역으로 나타났다. 특히 모든 지역에서 최소 조사구는 가구수가 1가구, 가구원수가 1명인 조사구가 존재하고, 최대 가구수는 지역에 따라 다소 차이를 보이고 있다. 최대 조사구를 지역

별로 보면, 서울에서는 2,120가구, 6,044명인 조사구가 존재하고, 대전은 1,235가구, 5,468명인 조사구가, 부산은 1,065가구, 3,184명인 조사구가, 경기도는 1,943가구, 2,678명인 조사구가 존재한다.

2) 병합조사구 현황

앞서 설명한 바와 같이 2015년 인구주택조사구는 기본적으로 평균 30가구 기준의 기본조사구로 구축되어 있다. 그러나 한국노동연구원과 협의 과정에서 2010년 인구주택조사구는 평균 60가구 기준으로 조사구를 구성하고 있으며, 또한 부재 및 주택의 고급화로 인해 표본가구의 접촉률 감소와 조사 참여 거부 증가 등으로 인한 조사 환경이 악화되어 1개 조사구 내에서 목표 패널 가구를 확보하기 위해 평균 30가구 기준의 기본조사구보다는 인접한 기본조사구를 병합한 평균 60가구 기준의 병합조사구를 구축하여 표본가구 추출 및 패널 가구를 확보하는 것이 바람직하다고 판단하였다. 그래서 통계청에서 추가로 제공하는 병합조사구를 표본추출틀로 사용하고자 병합조사구의 특성을 살펴보기로 한다. 병합조사구에 대한 가구수 및 가구원수(인구수) 특성을 간단히 살펴보면 <표 2-29>와 같다.

<표 2-29>를 보면, 평균 60가구인 2015년 인구주택조사구의 병합조사구수는 336,104개이다. 참고로, 2010년 인구주택조사구는 평균 60가구 기준의 조사구이며, 조사구수는 303,130개였다. 실제로 2015년 인구주택조사구에서 병합조사구의 평균 가구수는 56.8가구이고, 가구원수는 144.1명이다. 그러나 병합조사구의 가구수를 보면, 최소는 1개 가구, 최대 2,928가구를 갖는 조사구가 존재하며, 가구원수도 최소 1명, 최대 7,270명을 갖는 조사구가 있다. 병합조사구 내 가구수의 변동이 기본조사구보다 더 심한 것으로 나타났다.

<표 2-30>과 <표 2-31>은 병합조사구의 가구수와 가구원수 분포를 나타낸 결과이다. 조사구당 가구수가 40~60가구 미만인 조사구가 39.4%로 가장 많다. 가구수가 40~100가구인 조사구가 77.1%이고, 40가구 미만이 19.0%이다. 조사구당 가구원수가 120~160명 미만인 조사구가 26.4%로 가장 많다. 가구원수가 80~200명인 조사구가 64.4%이고, 80명 미만인 조사구가 17.5%이다.

〈표 2-29〉 2015년 인구주택조사구의 병합조사구 기준 가구수 및 인구수 특성

	N	평균값	표준편차	최소값	최대값
가구수	336,104	56.83	28.57	1	2,928
인구수	336,104	144.13	83.20	1	7,270

〈표 2-30〉 2015년 인구주택조사구의 병합조사구 기준 가구수 규모별 분포

	빈도	백분율	누적빈도	누적백분율
20가구 미만	18,076	5.38	18,076	5.38
20~40	45,892	13.65	63,968	19.03
40~60	132,360	39.38	196,328	58.41
60~80	95,473	28.41	291,801	86.82
80~100	31,439	9.35	323,240	96.17
100가구 이상	12,864	3.83	336,104	100.00

〈표 2-31〉 2015년 인구주택조사구의 병합조사구 기준 인구수 규모별 분포

	빈도	백분율	누적빈도	누적백분율
40명 미만	15,185	4.52	15,185	4.52
40~80	43,551	12.96	58,736	17.48
80~120	69,249	20.60	127,985	38.08
120~160	88,623	26.37	216,608	64.45
160~200	58,708	17.47	275,316	81.91
200~240	29,564	8.80	304,880	90.71
240~280	17,580	5.23	322,460	95.94
280명 이상	13,644	4.06	336,104	100.00

<표 2-32>와 <표 2-33>은 병합조사구의 가구수와 가구원수 규모별 가구수와 가구원수 특성을 요약한 결과이다. 가구수 규모별 최대 가구수는 100가구 이상인 규모에서 2,928가구인 조사구가 있으며, 최대 가구원수에서는 100가구 이상에서 최대 가구원수가 7,270명인 조사구가 있으며, 또한 40~60가구인 조사구에서 5,482명인 조사구가 존재하는 등 조사구의 가구수에 비해 가구원수가 많은 것으로 나타났다. 가구원수 규모별 가

〈표 2-32〉 2015년 인구주택조사구의 병합조사구 기준 가구수 규모별 가구수 및 인구수 특성

	조사구수	항목	평균	표준편차	최소값	최대값
20가구 미만	18,076	가구수	12.87	5.65	1	19
		인구수	32.26	20.39	1	985
20~40	45,892	가구수	30.43	5.80	20	39
		인구수	69.20	27.61	20	2,410
40~60	132,360	가구수	50.39	5.50	40	59
		인구수	126.86	37.37	40	5,482
60~80	95,473	가구수	67.84	5.75	60	79
		인구수	173.17	42.02	61	2,912
80~100	31,439	가구수	86.93	5.40	80	99
		인구수	231.66	51.13	83	1,112
100가구 이상	12,864	가구수	123.74	85.56	100	2,928
		인구수	316.34	219.91	105	7,270

〈표 2-33〉 2015년 인구주택조사구의 병합조사구 기준 인구수 규모별 가구수 및 인구수 특성

	조사구수	항목	평균	표준편차	최소값	최대값
40명 미만	15,185	가구수	13.83	7.94	1	39
		인구수	24.19	11.84	1	39
40~80	43,551	가구수	30.79	10.32	1	78
		인구수	60.38	11.35	40	79
80~120	69,249	가구수	47.90	10.11	2	116
		인구수	101.41	11.38	80	119
120~160	88,623	가구수	57.35	10.51	8	144
		인구수	139.19	11.39	120	159
160~200	58,708	가구수	65.50	12.01	11	180
		인구수	177.26	11.38	160	199
200~240	29,564	가구수	76.13	13.16	3	221
		인구수	217.64	11.60	200	239
240~280	17,580	가구수	85.63	12.87	22	232
		인구수	257.50	11.34	240	279
280명 이상	13,644	가구수	113.46	84.97	1	2,928
		인구수	345.38	214.75	280	7,270

〈표 2-34〉 2015년 인구주택조사구의 병합조사구 기준 광역시도별 가구수 및 인구수 특성

	조사구수	항목	평균	표준편차	최소값	최대값
서울	66,774	가구수	58.02	28.38	1	2,120
		인구수	143.69	75.78	1	6,044
부산	24,077	가구수	55.32	25.87	1	1,065
		인구수	137.95	72.49	1	3,184
대구	15,305	가구수	60.46	24.84	1	1,075
		인구수	155.00	76.71	1	2,906
인천	17,668	가구수	59.02	26.14	1	1,157
		인구수	156.44	79.04	1	3,076
광주	9,036	가구수	62.50	29.62	1	1,081
		인구수	158.73	79.67	1	2,561
대전	9,712	가구수	59.67	32.65	1	1,841
		인구수	150.90	113.49	1	5,482
울산	7,641	가구수	55.68	22.76	1	554
		인구수	145.33	70.40	1	1,329
세종	1,384	가구수	48.50	22.99	1	311
		인구수	124.14	66.14	1	481
경기	72,232	가구수	61.78	32.94	1	2,928
		인구수	165.28	97.28	1	7,270
강원	11,741	가구수	49.46	25.12	1	980
		인구수	117.42	69.39	1	2,514
충북	11,341	가구수	52.15	26.98	1	1,320
		인구수	129.49	76.42	1	3,138
충남	14,264	가구수	54.90	29.46	1	1,256
		인구수	134.96	79.05	1	2,422
전북	12,981	가구수	53.93	24.16	1	843
		인구수	131.14	68.42	1	1,607
전남	13,026	가구수	51.57	25.34	1	1,093
		인구수	123.08	69.48	1	2,532
경북	21,153	가구수	49.14	26.56	1	1,236
		인구수	117.38	80.10	1	3,169
경남	23,613	가구수	52.68	24.97	1	963
		인구수	132.77	76.89	1	2,588
제주	4,156	가구수	52.04	22.92	1	289
		인구수	136.37	62.42	1	659

구수를 보면 최소 1가구인 조사구가 존재하고, 또한 가구원수 규모별 가구수와 가구원수를 비교해 볼 때 가구원수와 가구수의 차이가 없는 것으로 보아 280명 미만 조사구에서는 1인 가구로 구성된 조사구가 존재하고 있음을 알 수 있다.

<표 2-34>는 광역시도별 병합조사구의 가구수와 가구원수 특성을 요약한 결과이다. 병합조사구수는 경기도가 72,232개로 가장 많고, 서울 66,774개, 부산 24,077개, 경남 23,613개, 경북 21,153개의 순으로 많다. 반면 세종시가 1,384개로 가장 적으며, 제주 4,156개, 울산 7,641개 순으로 적다. 광역시도별 평균 가구수는 광주가 62.5가구로 가장 많으며, 세종이 48.5가구로 가장 적다. 평균 가구원수는 경기가 165.3명으로 가장 많으며, 경북이 117.4명으로 가장 적다. 조사구별 변동은 경기가 32.9가구로 가장 심하고, 제주가 22.9가구로 가장 적다. 특이한 점은 광역시도별 최소 가구수가 1개인 조사구가 모두 존재하고, 최대 가구수는 지역별로 차이가 있으나 경기도, 서울에서는 2,000가구 이상인 조사구가 존재한다.

<표 2-35>는 조사구 특성별 병합조사구의 가구수와 가구원수 특성을 요약한 결과이다. 보통조사구가 183,276개이고, 아파트조사구가 152,828개이다. 평균 가구수는 보통조사구 53.8가구, 아파트조사구 60.4가구로 아파트조사구의 평균 가구수가 많으며, 평균 가구원수는 보통조사구 119.4명, 아파트조사구 173.7명으로 아파트조사구가 더 많다. 최소 가구수는 모두 1가구가 존재하고, 최대 가구수는 아파트조사구가 2,928가구, 보통조사구는 1,200가구인 조사구가 존재한다. 아파트조사구의 가구수 변동이 보통조사구보다 훨씬 심하고, 병합조사구의 가구 규모가 지나치게 많은 조사구가 존재하고 있다.

<표 2-36>은 동/읍면부별 병합조사구의 가구수와 가구원수 특성을 요약한 결과이다. 동부가 265,722개이고, 읍면부가 70,382개이다. 평균 가구수는 동부 58.7가구, 읍면부 49.8가구로 동부의 평균 가구수가 많으며, 평균 가구원수는 동부 150.5명, 읍면부 120.2명으로 동부가 더 많다. 최소 가구수는 모두 1가구가 존재하고, 최대 가구수는 읍면부가 2,928가구, 동부에서는 2,391가구인 조사구가 존재한다.

<표 2-37>은 조사구 특성별 동/읍면부 병합조사구의 가구수와 가구

〈표 2-35〉 2015년 인구주택조사구의 병합조사구 기준 조사구 특성별 가구수 및 인구수 특성

	조사구수	항목	평균	표준편차	최소값	최대값
보통(일반)	183,276	가구수	53.83	22.56	1	1,200
		인구수	119.43	58.15	1	5,482
아파트	152,828	가구수	60.42	34.08	1	2,928
		인구수	173.71	97.78	1	7,270

〈표 2-36〉 2015년 인구주택조사구의 병합조사구 기준 동/읍면부별 가구수 및 인구수 특성

	조사구수	항목	평균	표준편차	최소값	최대값
동부	265,722	가구수	58.70	28.34	1	2,391
		인구수	150.45	83.40	1	7,270
읍면부	70,382	가구수	49.77	28.36	1	2,928
		인구수	120.18	77.90	1	6,666

원수 특성을 요약한 결과이다. 동부는 아파트조사구가 134,213개이고, 보통조사구가 131,509개이다. 동부의 경우, 평균 가구수는 아파트조사구 61.0가구, 보통조사구 56.3가구로 동부 아파트조사구의 평균 가구수가 많으며, 평균 가구원수는 아파트조사구 176.0명, 보통조사구 124.4명으로 아파트조사구가 더 많다. 최소 가구수는 모두 1가구가 존재하고, 최대 가구수는 보통조사구가 1,200가구, 아파트조사구가 2,391가구인 조사구가 존재한다. 읍면부는 아파트조사구가 18,615개이고, 보통조사구가 51,767개이다. 읍면부의 경우, 평균 가구수는 아파트조사구 56.1가구, 보통조사구 47.5가구로 읍면부 아파트조사구의 평균 가구수가 많으며, 평균 가구원수는 아파트조사구 157.5명, 보통조사구 106.8명으로 아파트조사구가 더 많다. 최소 가구수는 모두 1가구가 존재하고, 최대 가구수는 보통조사구가 450가구, 아파트조사구가 2,928가구인 조사구가 존재한다.

<표 2-38>은 광역시도별 동읍면부 병합조사구의 가구수와 가구원수 특성을 요약한 결과이다. 광역시도의 동부와 읍면부의 평균 가구수 차이는 크지 않지만 경기도를 제외한 나머지 도지역은 동부와 읍면부의 평균

가구수 차이가 크며, 읍면부 병합조사구의 가구수 규모가 작다. 특히 경기도의 동부와 읍면부에서 조사구의 가구수가 2,391가구, 2,928가구인 조사구가 있는 것으로 나타났다.

〈표 2-37〉 2015년 인구주택조사구의 병합조사구 기준 동/읍면부 및 조사구 특성별 가구수 및 인구수 특성

	조사구특성	조사구수	항목	평균	표준편차	최소값	최대값
동부	보통(일반)	131,509	가구수	56.32	22.16	1	1,200
			인구수	124.41	57.19	1	5,482
	아파트	134,213	가구수	61.03	33.13	1	2,391
			인구수	175.96	96.19	1	7,270
읍면부	보통(일반)	51,767	가구수	47.50	22.31	1	450
			인구수	106.77	58.63	1	3,169
	아파트	18,615	가구수	56.06	40.04	1	2,928
			인구수	157.46	107.20	1	6,666

〈표 2-38〉 2015년 인구주택조사구의 병합조사구 기준 시도 및 동/읍면부별 가구수 및 인구수

		조사구수	항목	평균	표준편차	최소값	최대값
서울	동부	66,774	가구수	58.02	28.38	1	2,120
			인구수	143.69	75.78	1	6,044
부산	동부	22,829	가구수	56.00	24.58	1	772
			인구수	139.26	69.47	1	1,969
	읍면부	1,248	가구수	42.89	41.22	1	1,065
			인구수	113.89	111.78	3	3,184
대구	동부	14,170	가구수	60.59	24.88	1	1,075
			인구수	154.74	77.03	1	2,906
	읍면부	1,135	가구수	58.83	24.23	1	471
			인구수	158.26	72.57	1	1,089
인천	동부	17,178	가구수	59.14	26.27	1	1,157
			인구수	157.28	79.61	1	3,076
	읍면부	490	가구수	55.06	21.08	1	196
			인구수	127.27	47.24	2	306

〈표 2-38〉의 계속

		조사구수	항목	평균	표준편차	최소값	최대값
광주	동부	9,036	가구수	62.50	29.62	1	1,081
			인구수	158.73	79.67	1	2,561
대전	동부	9,712	가구수	59.67	32.65	1	1,841
			인구수	150.90	113.49	1	5,482
울산	동부	6,227	가구수	55.42	22.70	1	554
			인구수	144.77	70.31	1	1,329
	읍면부	1,414	가구수	56.80	22.98	2	325
			인구수	147.79	70.80	4	760
세종	동부	640	가구수	48.77	22.61	1	311
			인구수	138.11	67.58	1	481
	읍면부	744	가구수	48.27	23.33	2	270
			인구수	112.13	62.47	4	382
경기	동부	59,446	가구수	62.32	32.00	1	2,391
			인구수	167.48	97.02	1	7,270
	읍면부	12,786	가구수	59.24	36.87	1	2,928
			인구수	155.01	97.83	1	6,666
강원	동부	6,245	가구수	56.44	25.17	1	980
			인구수	137.20	72.49	1	2,514
	읍면부	5,496	가구수	41.54	22.58	1	326
			인구수	94.94	58.03	1	768
충북	동부	6,221	가구수	56.30	22.29	1	420
			인구수	144.24	69.11	1	1,085
	읍면부	5,120	가구수	47.11	31.02	1	1,320
			인구수	111.57	80.93	1	3,138
충남	동부	6,020	가구수	57.60	31.37	1	1,256
			인구수	146.08	81.97	1	1,737
	읍면부	8,244	가구수	52.93	27.81	1	918
			인구수	126.84	75.83	1	2,422
전북	동부	8,379	가구수	58.34	22.22	1	517
			인구수	146.32	68.51	1	1,236
	읍면부	4,602	가구수	45.91	25.45	1	843
			인구수	103.49	58.94	1	1,607

〈표 2-38〉의 계속

		조사구수	항목	평균	표준편차	최소값	최대값
전남	동부	5,030	가구수	57.97	29.96	1	1,093
			인구수	146.38	82.35	1	2,532
	읍면부	7,996	가구수	47.53	20.96	1	353
			인구수	108.42	55.16	1	683
경북	동부	10,180	가구수	53.68	28.01	1	1,236
			인구수	132.98	84.70	1	3,157
	읍면부	10,973	가구수	44.92	24.40	1	966
			인구수	102.92	72.65	1	3,169
경남	동부	14,518	가구수	55.71	25.19	1	963
			인구수	143.97	79.90	1	2,588
	읍면부	9,095	가구수	47.85	23.82	1	734
			인구수	114.89	68.12	1	2,000
제주	동부	3,117	가구수	51.29	23.27	1	289
			인구수	135.73	64.13	1	659
	읍면부	1,039	가구수	54.29	21.70	1	162
			인구수	138.30	56.96	3	376

<표 2-39>는 광역시도별, 조사구 특성별 병합조사구의 가구수와 가구원수 특성을 요약한 결과이다. 서울, 대구, 대전, 강원, 충북, 충남, 전북, 전남, 경북, 경남, 제주 지역은 보통조사구가 아파트조사구보다 더 많다.

〈표 2-39〉 2015년 인구주택조사구의 병합조사구 기준 시도 및 조사구 특성별 가구수 및 인구수

	조사구특성	조사구수	항목	평균	표준편차	최소값	최대값
서울	보통 (일반)	38,841	가구수	58.25	23.77	1	1,200
			인구수	128.04	52.89	1	1,846
	아파트	27,933	가구수	57.70	33.76	1	2,120
			인구수	165.44	95.01	1	6,044
부산	보통 (일반)	11,315	가구수	55.87	21.78	1	617
			인구수	121.78	51.47	1	817
	아파트	12,762	가구수	54.82	29.01	1	1,065
			인구수	152.28	84.43	1	3,184

〈표 2-39〉의 계속

	조사구특성	조사구수	항목	평균	표준편차	최소값	최대값
대구	보통(일반)	7,998	가구수	54.79	17.37	1	444
			인구수	122.55	43.26	1	749
	아파트	7,307	가구수	66.67	29.80	1	1,075
			인구수	190.51	88.69	1	2,906
인천	보통(일반)	8,523	가구수	56.10	20.46	1	425
			인구수	130.36	55.15	1	948
	아파트	9,145	가구수	61.75	30.25	1	1,157
			인구수	180.76	89.50	2	3,076
광주	보통(일반)	3,559	가구수	58.54	30.52	1	1,081
			인구수	121.34	57.34	1	1,544
	아파트	5,477	가구수	65.08	28.74	1	967
			인구수	183.03	82.69	2	2,561
대전	보통(일반)	4,925	가구수	53.16	21.39	1	352
			인구수	113.37	93.74	1	5,482
	아파트	4,787	가구수	66.36	40.04	1	1,841
			인구수	189.52	118.96	1	5,330
울산	보통(일반)	3,345	가구수	55.54	17.50	1	179
			인구수	121.49	43.51	1	382
	아파트	4,296	가구수	55.78	26.13	1	554
			인구수	163.90	80.96	1	1,329
세종	보통(일반)	523	가구수	47.10	23.59	2	270
			인구수	97.51	53.35	4	345
	아파트	861	가구수	49.36	22.60	1	311
			인구수	140.32	67.93	1	481
경기	보통(일반)	35,543	가구수	57.26	21.96	1	793
			인구수	131.99	61.07	1	2,091
	아파트	36,689	가구수	66.16	40.37	1	2,928
			인구수	197.52	113.60	1	7,270
강원	보통(일반)	7,271	가구수	44.70	22.23	1	200
			인구수	95.98	51.81	1	637

〈표 2-39〉의 계속

	조사구특성	조사구수	항목	평균	표준편차	최소값	최대값
강원	아파트	4,470	가구수	57.20	27.50	1	980
			인구수	152.30	79.49	1	2,514
충북	보통(일반)	6,680	가구수	47.89	22.55	1	290
			인구수	106.39	54.84	1	694
	아파트	4,661	가구수	58.25	31.28	1	1,320
			인구수	162.59	89.68	1	3,138
충남	보통(일반)	8,677	가구수	51.58	25.26	1	718
			인구수	112.96	58.79	1	1,039
	아파트	5,587	가구수	60.06	34.36	1	1,256
			인구수	169.14	93.09	1	2,422
전북	보통(일반)	7,630	가구수	49.35	21.93	1	447
			인구수	105.98	49.10	1	561
	아파트	5,351	가구수	60.48	25.65	1	843
			인구수	167.01	75.69	1	1,607
전남	보통(일반)	8,507	가구수	48.51	20.73	1	267
			인구수	104.89	48.84	1	683
	아파트	4,519	가구수	57.32	31.49	1	1,093
			인구수	157.31	87.36	2	2,532
경북	보통(일반)	13,950	가구수	45.99	20.76	1	272
			인구수	98.53	63.92	1	3,169
	아파트	7,203	가구수	55.24	34.36	1	1,236
			인구수	153.91	94.38	1	3,157
경남	보통(일반)	13,173	가구수	49.87	19.63	1	383
			인구수	109.06	56.53	1	2,410
	아파트	10,440	가구수	56.23	30.02	1	963
			인구수	162.69	87.96	1	2,588
제주	보통(일반)	2,816	가구수	56.56	21.11	1	289
			인구수	144.02	57.92	1	656
	아파트	1,340	가구수	42.54	23.67	1	281
			인구수	120.30	68.22	1	659

3. 노동패널 현황

2017통합패널을 위한 새로운 표본설계에서는 2015년 인구주택총조사를 모집단 및 조사구로 선정하여 광역시도별 동읍면부 및 주택유형별 표본을 배분한 후, 1998년 및 2009년 패널에 대한 연속성을 유지하도록 패널을 구축해야 하므로, 2017통합패널에서는 층(광역시도별 동읍면부 및 주택유형)별로 현행 패널을 우선 확정하고 층별로 부족한 패널을 추가로 추출하고자 한다. 따라서 본항에서는 한국노동패널의 19차 조사 패널 현황을 살펴보고, 2015년 인구주택총조사 특성과 비교하고자 한다.

가. 19차 노동패널 현황

한국노동패널 19차 조사 결과의 광역시도별 동읍면부 및 주택유형별 가구수 현황은 <표 2-40>과 같다. 한국노동패널 19차 조사에 응답 완료한 가구수는 7,012가구이다. 동읍면부별로는 동부가 5,673가구이고, 읍면부는 1,339가구이다. 주택유형별로는 아파트조사구가 3,397가구, 보통조사구가 3,615가구이다. 보통조사구 중 단독주택이 1,814가구이고, 연립/다세대/다가구 등이 1,801가구이다. 동부에서는 아파트조사구가 2,989가구이고, 보통조사구가 2,684가구이며, 읍면부에서는 아파트조사구가 408가구이고, 보통조사구가 931가구이다.

〈표 2-40〉 19차 노동패널의 광역시도별, 동/읍면부별 및 주택유형별 가구수

		가구수	아파트	보통조사구						
				소계	단독	단독 이외				
						소계	연립	다세대	상가	기타
전국	계	7,012	3,397	3,615	1,814	1,801	680	786	230	105
	동부	5,673	2,989	2,684	1,078	1,606	616	717	185	88
	읍면부	1,339	408	931	736	195	64	69	45	17
서울	소계	1,297	553	744	209	535	219	257	37	22
	동부	1,297	553	744	209	535	219	257	37	22
	읍면부									

〈표 2-40〉의 계속

		가구수	아파트	보통조사구						
				소계	단독	단독 이외				
						소계	연립	다세대	상가	기타
부산	소계	526	250	276	142	134	50	63	12	9
	동부	507	238	269	139	130	48	62	12	8
	읍면부	19	12	7	3	4	2	1		1
대구	소계	327	172	155	89	66	12	43	6	5
	동부	303	164	139	77	62	10	41	6	5
	읍면부	24	8	16	12	4	2	2		
인천	소계	413	224	189	63	126	91	14	16	5
	동부	400	223	177	53	124	91	13	16	4
	읍면부	13	1	12	10	2		1		1
광주	소계	186	116	70	48	22	4	10	6	2
	동부	186	116	70	48	22	4	10	6	2
	읍면부									
대전	소계	186	97	89	27	62	19	30	10	3
	동부	186	97	89	27	62	19	30	10	3
	읍면부									
울산	소계	190	111	79	29	50	14	22	9	5
	동부	155	90	65	23	42	12	18	7	5
	읍면부	35	21	14	6	8	2	4	2	
세종	소계	20	13	7	5	2		2		
	동부	11	11	0		0				
	읍면부	9	2	7	5	2		2		
경기	소계	1,534	897	637	205	432	173	202	33	24
	동부	1,309	783	526	143	383	149	182	30	22
	읍면부	225	114	111	62	49	24	20	3	2
강원	소계	195	93	102	81	21	8	9	4	
	동부	130	84	46	31	15	4	7	4	
	읍면부	65	9	56	50	6	4	2		
충북	소계	247	108	139	90	49	13	20	13	3
	동부	147	83	64	34	30	8	11	9	2
	읍면부	100	25	75	56	19	5	9	4	1

〈표 2-40〉의 계속

		가구수	아파트	보통조사구						
				소계	단독	단독 이외				
						소계	연립	다세대	상가	기타
충남	소계	338	123	215	155	60	21	20	13	6
	동부	145	73	72	37	35	10	16	7	2
	읍면부	193	50	143	118	25	11	4	6	4
전북	소계	265	104	161	132	29	9	6	8	6
	동부	176	95	81	56	25	9	6	6	4
	읍면부	89	9	80	76	4			2	2
전남	소계	296	105	191	155	36	3	10	19	4
	동부	127	77	50	30	20		7	10	3
	읍면부	169	28	141	125	16	3	3	9	1
경북	소계	408	162	246	156	90	23	38	24	5
	동부	208	104	104	52	52	17	23	10	2
	읍면부	200	58	142	104	38	6	15	14	3
경남	소계	521	258	263	185	78	15	38	19	6
	동부	338	187	151	90	61	11	32	14	4
	읍면부	183	71	112	95	17	4	6	5	2
제주	소계	63	11	52	43	9	6	2	1	
	동부	48	11	37	29	8	5	2	1	
	읍면부	15		15	14	1	1			

<표 2-41>은 19차 노동패널의 광역시도별 동읍면부 및 가구원수별 패널 현황이다. 전체 패널의 평균 가구원수는 2.67명이고, 동부가 2.72명, 읍면부가 2.48명이다. 가구원수별 분포를 보면, 2인 가구가 24.9%로 가장 많으며, 4인 가구가 24.7%, 1인 가구가 23.7%의 순으로 많다. 이는 98패널과 09통합패널 구축 시점의 일반가구의 가구원수가 2인 가구 혹은 4인 가구가 대표 가구이므로 현재 노동패널의 가구원수는 2인 혹은 4인 비중이 많은 것으로 보인다.

동부의 경우는 4인 가구, 2인 가구, 1인 가구, 3인 가구 순으로 많고, 읍면부는 2인 가구, 1인 가구, 4인 가구, 3인 가구의 순으로 많다. 광역시와

〈표 2-41〉 광역시도별, 동/읍면부별 및 가구원수별 가구수

		가구수	가구원수										
			1		2		3		4		5+		평균
			가구수	%	가구수	%	가구수	%	가구수	%	가구수	%	
전국	소계	7,012	1,659	23.66	1,744	24.87	1,415	20.18	1,732	24.70	462	6.59	2.67
	동부	5,673	1,283	22.62	1,326	23.37	1,213	21.38	1,489	26.25	362	6.38	2.72
	읍면부	1,339	376	28.08	418	31.22	202	15.09	243	18.15	100	7.47	2.48
서울	소계	1,297	307	23.67	313	24.13	306	23.59	301	23.21	70	5.40	2.64
	동부	1,297	307	23.67	313	24.13	306	23.59	301	23.21	70	5.40	2.64
	읍면부												
부산	소계	526	135	25.67	134	25.48	113	21.48	122	23.19	22	4.18	2.56
	동부	507	132	26.04	127	25.05	112	22.09	116	22.88	20	3.94	2.55
	읍면부	19	3	15.79	7	36.84	1	5.26	6	31.58	2	10.53	2.84
대구	소계	327	84	25.69	76	23.24	64	19.57	76	23.24	27	8.26	2.66
	동부	303	77	25.41	71	23.43	59	19.47	71	23.43	25	8.25	2.67
	읍면부	24	7	29.17	5	20.83	5	20.83	5	20.83	2	8.33	2.58
인천	소계	413	76	18.40	96	23.24	90	21.79	118	28.57	33	7.99	2.86
	동부	400	74	18.50	91	22.75	90	22.50	114	28.50	31	7.75	2.86
	읍면부	13	2	15.38	5	38.46			4	30.77	2	15.38	2.92
광주	소계	186	52	27.96	52	27.96	29	15.59	40	21.51	13	6.99	2.53
	동부	186	52	27.96	52	27.96	29	15.59	40	21.51	13	6.99	2.53
	읍면부												
대전	소계	186	44	23.66	42	22.58	43	23.12	49	26.34	8	4.30	2.65
	동부	186	44	23.66	42	22.58	43	23.12	49	26.34	8	4.30	2.65
	읍면부												
울산	소계	190	35	18.42	33	17.37	48	25.26	63	33.16	11	5.79	2.92
	동부	155	28	18.06	26	16.77	39	25.16	54	34.84	8	5.16	2.94
	읍면부	35	7	20.00	7	20.00	9	25.71	9	25.71	3	8.57	2.83
세종	소계	20	4	20.00	4	20.00	6	30.00	5	25.00	1	5.00	2.75
	동부	11	1	9.09			6	54.55	3	27.27	1	9.09	3.27
	읍면부	9	3	33.33	4	44.44			2	22.22			2.11
경기	소계	1,534	279	18.19	333	21.71	333	21.71	472	30.77	117	7.63	2.90
	동부	1,309	238	18.18	285	21.77	290	22.15	405	30.94	91	6.95	2.88
	읍면부	225	41	18.22	48	21.33	43	19.11	67	29.78	26	11.56	3.00

〈표 2-41〉의 계속

		가구수	가구원수										
			1		2		3		4		5+		평균
			가구수	%	가구수	%	가구수	%	가구수	%	가구수	%	
강원	소계	195	46	23.59	57	29.23	32	16.41	42	21.54	18	9.23	2.66
	동부	130	28	21.54	34	26.15	22	16.92	33	25.38	13	10.00	2.80
	읍면부	65	18	27.69	23	35.38	10	15.38	9	13.85	5	7.69	2.38
충북	소계	247	72	29.15	64	25.91	41	16.60	61	24.70	9	3.64	2.49
	동부	147	36	24.49	34	23.13	24	16.33	46	31.29	7	4.76	2.71
	읍면부	100	36	36.00	30	30.00	17	17.00	15	15.00	2	2.00	2.17
충남	소계	338	102	30.18	85	25.15	51	15.09	66	19.53	34	10.06	2.57
	동부	145	41	28.28	31	21.38	28	19.31	31	21.38	14	9.66	2.65
	읍면부	193	61	31.61	54	27.98	23	11.92	35	18.13	20	10.36	2.52
전북	소계	265	76	28.68	73	27.55	44	16.60	52	19.62	20	7.55	2.52
	동부	176	43	24.43	46	26.14	30	17.05	44	25.00	13	7.39	2.66
	읍면부	89	33	37.08	27	30.34	14	15.73	8	8.99	7	7.87	2.25
전남	소계	296	85	28.72	101	34.12	45	15.20	53	17.91	12	4.05	2.35
	동부	127	31	24.41	31	24.41	25	19.69	33	25.98	7	5.51	2.65
	읍면부	169	54	31.95	70	41.42	20	11.83	20	11.83	5	2.96	2.12
경북	소계	408	115	28.19	125	30.64	72	17.65	75	18.38	21	5.15	2.44
	동부	208	56	26.92	52	25.00	44	21.15	46	22.12	10	4.81	2.54
	읍면부	200	59	29.50	73	36.50	28	14.00	29	14.50	11	5.50	2.34
경남	소계	521	134	25.72	133	25.53	91	17.47	125	23.99	38	7.29	2.62
	동부	338	86	25.44	73	21.60	61	18.05	93	27.51	25	7.40	2.70
	읍면부	183	48	26.23	60	32.79	30	16.39	32	17.49	13	7.10	2.48
제주	소계	63	13	20.63	23	36.51	7	11.11	12	19.05	8	12.70	2.68
	동부	48	9	18.75	18	37.50	5	10.42	10	20.83	6	12.50	2.71
	읍면부	15	4	26.67	5	33.33	2	13.33	2	13.33	2	13.33	2.60

경기도를 제외한 도지역에서 2인 가구 및 1인 가구의 패널이 많다. 특이한 점은 전남의 읍면부에서 2인 가구가 41.4%이다.

나. 모집단과 19차 노동패널의 구성비 비교

새로운 표본설계의 목적은 17통합패널을 위한 표본설계 방안을 마련하는 것이다. 17통합패널은 현행 19차 노동패널과 2015년 모집단으로부터 새롭게 추출한 추가 패널로 구축하고자 한다.

본항에서는 17통합패널 구축에 필요한 새로운 추가 패널의 규모 및 층별 배분을 위해 2015년 인구주택총조사와 한국노동패널의 19차 패널을 비교해 추가 패널 규모 및 배분 방안을 검토해 보고자 한다.

<표 2-42>는 광역시도별 동읍면부 및 주택유형별 가구수의 구성비를 비교한 결과이다. 2015년 모집단에서 동부 및 읍면부는 81.0%, 19.0%이고, 19차 노동패널은 80.9%, 19.1%이며, 주택유형별로는 모집단에서 아파트가 48.1%, 주택이 51.9%이고, 현행 노동패널에서는 아파트가 48.5%, 주택이 51.5%로 나타났다. 그러나 광역시도 내 아파트 및 주택 구성비는 모집단과 노동패널에 약간의 차이가 있고, 또한 일반조사구 내 단독 주택과 연립/다세대/다가구의 구성비는 동부에서 노동패널의 구성비가 모집단 구성비보다 높게 나타났다.

<표 2-43>은 광역시도별 동읍면부 및 가구원수별 가구수의 구성비를

〈표 2-42〉 광역시도별, 동/읍면부별 및 주택유형별 가구수 구성비 비교

		모집단					표본(노동패널 19차)				
		가구수 비율	아파트	일반조사구			가구수 비율	아파트	일반조사구		
				소계	단독	이외[1]			소계	단독	이외[1]
전국	소계	100.00	48.12	51.88	35.26	16.62	100.00	48.45	51.55	25.87	25.68
	동부	81.04	52.61	47.39	29.64	17.75	80.90	52.69	47.31	19.00	28.31
	읍면부	18.96	28.94	71.06	59.29	11.77	19.10	30.47	69.53	54.97	14.56
서울	소계	19.80	41.92	58.08	31.38	26.71	18.50	42.64	57.36	16.11	41.25
	동부	19.80	41.92	58.08	31.38	26.71	18.50	42.64	57.36	16.11	41.25
	읍면부										
부산	소계	6.99	51.77	48.23	30.96	17.27	7.50	47.53	52.47	27.00	25.48
	동부	6.71	51.13	48.87	31.37	17.51	7.23	46.94	53.06	27.42	25.64
	읍면부	0.28	67.06	32.94	21.26	11.68	0.27	63.16	36.84	15.79	21.05

〈표 2-42〉의 계속

		모집단					표본(노동패널 19차)				
		가구수 비율	아파트	일반조사구			가구수 비율	아파트	일반조사구		
				소계	단독	이외[1]			소계	단독	이외[1]
대구	소계	4.86	52.41	47.59	37.68	9.90	4.66	52.60	47.40	27.22	20.18
	동부	4.51	51.54	48.46	38.44	10.02	4.32	54.13	45.87	25.41	20.46
	읍면부	0.35	63.73	36.27	27.84	8.44	0.34	33.33	66.67	50.00	16.67
인천	소계	5.47	53.12	46.88	19.54	27.34	5.89	54.24	45.76	15.25	30.51
	동부	5.29	54.68	45.32	17.70	27.61	5.70	55.75	44.25	13.25	31.00
	읍면부	0.18	6.77	93.23	73.91	19.31	0.19	7.69	92.31	76.92	15.38
광주	소계	2.97	62.72	37.28	31.10	6.18	2.65	62.37	37.63	25.81	11.83
	동부	2.97	62.72	37.28	31.10	6.18	2.65	62.37	37.63	25.81	11.83
	읍면부										
대전	소계	3.05	54.60	45.40	35.10	10.30	2.65	52.15	47.85	14.52	33.33
	동부	3.05	54.60	45.40	35.10	10.30	2.65	52.15	47.85	14.52	33.33
	읍면부										
울산	소계	2.22	56.15	43.85	33.39	10.46	2.71	58.42	41.58	15.26	26.32
	동부	1.81	56.64	43.36	32.71	10.65	2.21	58.06	41.94	14.84	27.10
	읍면부	0.41	53.95	46.05	36.41	9.64	0.50	60.00	40.00	17.14	22.86
세종	소계	0.39	62.38	37.62	31.06	6.56	0.29	65.00	35.00	25.00	10.00
	동부	0.18	97.00	3.00	0.17	2.83	0.16	100.00	0.00	0.00	0.00
	읍면부	0.21	32.38	67.62	57.83	9.79	0.13	22.22	77.78	55.56	22.22
경기	소계	22.94	54.97	45.03	25.99	19.04	21.88	58.47	41.53	13.36	28.16
	동부	19.05	58.13	41.87	23.22	18.66	18.67	59.82	40.18	10.92	29.26
	읍면부	3.90	39.55	60.45	39.53	20.92	3.21	50.67	49.33	27.56	21.78
강원	소계	3.17	43.10	56.90	47.86	9.04	2.78	47.69	52.31	41.54	10.77
	동부	1.89	57.14	42.86	35.80	7.06	1.85	64.62	35.38	23.85	11.54
	읍면부	1.28	22.33	77.67	65.70	11.97	0.93	13.85	86.15	76.92	9.23
충북	소계	3.15	45.53	54.47	45.41	9.06	3.52	43.72	56.28	36.44	19.84
	동부	1.85	57.62	42.38	34.67	7.70	2.10	56.46	43.54	23.13	20.41
	읍면부	1.30	28.41	71.59	60.60	10.98	1.43	25.00	75.00	56.00	19.00
충남	소계	4.17	42.58	57.42	46.59	10.83	4.82	36.39	63.61	45.86	17.75
	동부	1.83	59.50	40.50	28.83	11.67	2.07	50.34	49.66	25.52	24.14
	읍면부	2.34	29.35	70.65	60.47	10.18	2.75	25.91	74.09	61.14	12.95

〈표 2-42〉의 계속

		모집단					표본(노동패널 19차)				
		가구수 비율	아파트	일반조사구			가구수 비율	아파트	일반조사구		
				소계	단독	이외[1]			소계	단독	이외[1]
전북	소계	3.75	45.60	54.40	47.54	6.85	3.78	39.25	60.75	49.81	10.94
	동부	2.58	59.85	40.15	33.26	6.90	2.51	53.98	46.02	31.82	14.20
	읍면부	1.17	14.21	85.79	79.02	6.77	1.27	10.11	89.89	85.39	4.49
전남	소계	3.77	36.30	63.70	55.91	7.79	4.22	35.47	64.53	52.36	12.16
	동부	1.55	61.52	38.48	30.06	8.43	1.81	60.63	39.37	23.62	15.75
	읍면부	2.22	18.61	81.39	74.05	7.35	2.41	16.57	83.43	73.96	9.47
경북	소계	5.56	37.79	62.21	51.86	10.36	5.82	39.71	60.29	38.24	22.06
	동부	2.89	52.76	47.24	35.83	11.41	2.97	50.00	50.00	25.00	25.00
	읍면부	2.67	21.53	78.47	69.26	9.21	2.85	29.00	71.00	52.00	19.00
경남	소계	6.59	46.70	53.30	44.34	8.96	7.43	49.52	50.48	35.51	14.97
	동부	4.24	55.44	44.56	34.97	9.59	4.82	55.33	44.67	26.63	18.05
	읍면부	2.35	30.91	69.09	61.25	7.84	2.61	38.80	61.20	51.91	9.29
제주	소계	1.15	25.36	74.64	50.53	24.11	0.90	17.46	82.54	68.25	14.29
	동부	0.84	33.25	66.75	40.12	26.63	0.68	22.92	77.08	60.42	16.67
	읍면부	0.31	3.91	96.09	78.82	17.27	0.21	0.00	100.00	93.33	6.67

주 : 1) 이외는 연립/다세대/다가구 주택임.

〈표 2-43〉 광역시도별, 동/읍면부별 및 가구원수별 가구수 구성비 비교

	모집단						표본(노동패널 19차)					
	전체	1명	2명	3명	4명	5명	전체	1명	2명	3명	4명	5명
전국	100.00	27.23	26.13	21.46	18.78	6.40	100.00	23.66	24.87	20.18	24.70	6.59
동부	81.04	26.63	24.82	22.15	19.97	6.42	80.90	22.62	23.37	21.38	26.25	6.38
읍면부	18.96	29.77	31.72	18.50	13.70	6.32	19.10	28.08	31.22	15.09	18.15	7.47
서울	19.80	29.48	24.59	21.60	18.55	5.78	18.50	23.67	24.13	23.59	23.21	5.40
동부	19.80	29.48	24.59	21.60	18.55	5.78	18.50	23.67	24.13	23.59	23.21	5.40
읍면부												
부산	6.99	27.08	27.40	22.21	17.82	5.49	7.50	25.67	25.48	21.48	23.19	4.18
동부	7.07	25.89	25.96	21.02	16.80	5.16	7.23	26.04	25.05	22.09	22.88	3.94
읍면부	0.30	20.54	26.26	21.99	18.93	6.14	0.27	15.79	36.84	5.26	31.58	10.53

〈표 2-43〉의 계속

	모집단						표본(노동패널 19차)					
	전체	1명	2명	3명	4명	5명	전체	1명	2명	3명	4명	5명
대구	4.86	25.80	25.83	22.49	19.84	6.05	4.66	25.69	23.24	19.57	23.24	8.26
동부	4.78	24.65	24.32	21.17	18.57	5.65	4.32	25.41	23.43	19.47	23.43	8.25
읍면부	0.37	20.08	24.88	21.70	20.51	6.42	0.34	29.17	20.83	20.83	20.83	8.33
인천	5.47	23.31	25.36	23.45	21.10	6.79	5.89	18.40	23.24	21.79	28.57	7.99
동부	5.29	23.01	25.00	23.68	21.47	6.84	5.70	18.50	22.75	22.50	28.50	7.75
읍면부	0.18	31.65	35.39	16.45	9.71	5.25	0.19	15.38	38.46	0.00	30.77	15.38
광주	2.97	28.84	24.27	20.40	19.33	7.16	2.65	27.96	27.96	15.59	21.51	6.99
동부	2.97	28.84	24.27	20.40	19.33	7.16	2.65	27.96	27.96	15.59	21.51	6.99
읍면부												
대전	3.05	29.08	24.14	20.96	19.24	6.59	2.65	23.66	22.58	23.12	26.34	4.30
동부	3.05	29.08	24.14	20.96	19.24	6.59	2.65	23.66	22.58	23.12	26.34	4.30
읍면부												
울산	2.22	24.46	24.59	23.67	21.43	5.86	2.71	18.42	17.37	25.26	33.16	5.79
동부	1.81	24.49	24.23	23.94	21.58	5.76	2.21	18.06	16.77	25.16	34.84	5.16
읍면부	0.43	23.25	25.03	21.53	19.89	6.00	0.50	20.00	20.00	25.71	25.71	8.57
세종	0.39	29.10	23.94	20.23	19.36	7.37	0.29	20.00	20.00	30.00	25.00	5.00
동부	0.18	20.33	20.80	24.01	26.93	7.93	0.16	9.09	0.00	54.55	27.27	9.09
읍면부	0.22	35.42	25.74	16.37	12.35	6.64	0.13	33.33	44.44	0.00	22.22	0.00
경기	22.94	23.41	24.23	22.93	22.15	7.29	21.88	18.19	21.71	21.71	30.77	7.63
동부	19.05	22.99	23.39	23.38	23.03	7.21	18.67	18.18	21.77	22.15	30.94	6.95
읍면부	3.90	25.44	28.33	20.71	17.83	7.69	3.21	18.22	21.33	19.11	29.78	11.56
강원	3.17	31.24	29.44	19.20	14.38	5.73	2.78	23.59	29.23	16.41	21.54	9.23
동부	1.89	29.87	27.52	20.43	16.43	5.75	1.85	21.54	26.15	16.92	25.38	10.00
읍면부	1.28	33.28	32.30	17.38	11.35	5.69	0.93	27.69	35.38	15.38	13.85	7.69
충북	3.15	28.84	27.77	20.04	16.70	6.64	3.52	29.15	25.91	16.60	24.70	3.64
동부	1.85	27.01	25.05	21.58	19.41	6.96	2.10	24.49	23.13	16.33	31.29	4.76
읍면부	1.30	31.44	31.62	17.87	12.88	6.19	1.43	36.00	30.00	17.00	15.00	2.00
충남	4.17	29.45	28.47	19.29	16.12	6.66	4.82	30.18	25.15	15.09	19.53	10.06
동부	1.83	28.99	23.57	20.71	19.86	6.87	2.07	28.28	21.38	19.31	21.38	9.66
읍면부	2.34	29.82	32.30	18.18	13.20	6.50	2.75	31.61	27.98	11.92	18.13	10.36

〈표 2-43〉의 계속

	모집단						표본(노동패널 19차)					
	전체	1명	2명	3명	4명	5명	전체	1명	2명	3명	4명	5명
전북	3.75	29.80	29.14	18.94	15.35	6.77	3.78	28.68	27.55	16.60	19.62	7.55
동부	2.58	28.66	26.17	20.17	17.91	7.09	2.51	24.43	26.14	17.05	25.00	7.39
읍면부	1.17	32.31	35.68	16.23	9.72	6.06	1.27	37.08	30.34	15.73	8.99	7.87
전남	3.77	30.37	31.64	18.24	13.45	6.30	4.22	28.72	34.12	15.20	17.91	4.05
동부	1.55	26.92	27.78	20.86	17.55	6.88	1.81	24.41	24.41	19.69	25.98	5.51
읍면부	2.22	32.79	34.34	16.40	10.58	5.89	2.41	31.95	41.42	11.83	11.83	2.96
경북	5.56	30.35	30.16	19.52	14.71	5.26	5.82	28.19	30.64	17.65	18.38	5.15
동부	2.89	28.73	26.52	21.32	17.88	5.54	2.97	26.92	25.00	21.15	22.12	4.81
읍면부	2.67	32.11	34.11	17.55	11.27	4.95	2.85	29.50	36.50	14.00	14.50	5.50
경남	6.59	27.55	27.16	21.07	18.13	6.09	7.43	25.72	25.53	17.47	23.99	7.29
동부	4.24	26.24	24.62	22.44	20.47	6.23	4.82	25.44	21.60	18.05	27.51	7.40
읍면부	2.35	29.92	31.72	18.60	13.92	5.84	2.61	26.23	32.79	16.39	17.49	7.10
제주	1.15	26.52	26.46	19.90	17.38	9.73	0.90	20.63	36.51	11.11	19.05	12.70
동부	0.84	26.23	24.77	20.39	18.87	9.73	0.68	18.75	37.50	10.42	20.83	12.50
읍면부	0.31	27.31	31.03	18.58	13.35	9.73	0.21	26.67	33.33	13.33	13.33	13.33

비교한 결과이다. 2015년 모집단과 19차 노동패널의 가구원수별 가구수 구성비를 비교해 보면, 1인 가구와 2인 가구는 노동패널이 3.6%, 1.3% 작지만 4인 가구는 5.9% 더 높다. 특히 동부에서 이러한 경향이 나타나고 있다. 광역시도별로 가구원수 구성비를 보면, 세종시는 98패널 및 09통합패널에서는 충남과 충북에 포함된 지역이므로 현행 노동패널과 모집단의 가구원수 구성비에 차이가 다소 크게 나타나고 있다.

<표 2-42>와 <표 2-43>에서 17통합패널 구축을 위해 2015년 모집단과 19차 노동패널의 광역시도별, 동읍면부별 주택유형과 가구원수 구성비를 비교해 본 결과, 가구원수 구성비가 약간 차이가 있지만 전반적으로 현재 19차 노동패널의 광역시도별 동읍면부 및 주택유형별 구성비는 2015년 모집단 구성비와 큰 차이가 없는 것으로 나타났다. 이는 98패널과 09통합패널을 기준으로 구축한 현행 노동패널 내 패널 가구의 변화가 모

집단 가구 특성의 변화와 동일한 경향을 보이기 때문으로 생각된다.

4. 시범조사를 위한 새로운 표본 설계

2017통합패널을 위한 새로운 표본설계는 광역시도별 동읍면부 및 주택유형별로 1998년 및 2009년에 구축되어 현재까지 지속적으로 참여하는 패널을 포함하기 때문에 2017통합패널에서는 광역시도별 동읍면부 및 주택유형(층)별로 현행 패널을 우선 확정하고 층별로 부족한 패널을 추가로 추출하고자 한다. 즉 17통합패널은 기본적으로 1998년 혹은 2009년 구축된 패널 구축을 위한 표본설계의 틀을 유지하고, 2015년 인구주택총조사에서 나타난 광역시도별 동읍면부 및 주택유형별 변화된 모집단 특성과 현행 패널에서 추적 실패 및 패널 탈락 등으로 인해 마모된 패널을 보완하도록 표본설계를 수행한다.

가. 표본크기 결정 및 배분 시 고려 사항

2017통합패널을 위한 새로운 표본설계에서는 다음과 같은 사항을 고려해 표본설계를 수행한다.

첫째, 98년 및 09통합패널 구축을 위한 표본설계의 틀을 동일하게 유지한다.

둘째, 전국 및 광역시도별 분석을 위한 최소 표본규모를 확보하기 위해 세종과 제주도의 표본가구수를 고려한다.

셋째, 광역시도별 동읍면부 및 주택유형별 표본은 다음의 3단계로 배분한다.

- 1단계 : 광역시도
- 2단계 : 동/읍면부
- 3단계 : 조사구 특성(주택유형과 동일하며, 보통 및 아파트조사구로 층화)

넷째, 조사구는 기본조사구를 병합한 평균 60가구 기준의 병합조사구를 사용하며, 병합조사구당 10가구를 추출하도록 조사구수를 결정한다.

나. 17통합패널 구축을 위한 새로운 표본설계

17통합패널을 위한 새로운 표본설계에서는 통합패널의 크기에 대한 적절성을 검토하고, 2015년 모집단에서 추가로 추출할 표본크기 결정 및 층별 배분 방안을 제시한다.

1) 17통합패널의 표본 규모의 적절성 검토

17통합패널의 목표 및 조사모집단은 섬조사구 및 특수조사구를 제외한 보통조사구 및 아파트조사구에 거주하는 일반가구이다. 표본추출을 위한 표본추출틀은 2015년 인구주택총조사로 구축한 기본조사구를 병합한 평균 60가구 기준의 병합조사구이다.

17통합패널은 19차 패널과 새롭게 추출한 추가패널을 추가하여 구축한다. 현행 패널에 추가하여 통합패널을 구축하려는 주된 이유는 ① 98년과 09년 패널에 대한 종단 연구를 위해 지속적인 자료를 제공할 필요가 있으며, ② 노동패널조사에서 제공하는 노동 환경의 변화로 인해 변화된 노동시장에서의 심층적인 분석 요구가 증가하고, 가구 및 가구원 특성을 고려한 다양한 노동실태 파악을 위한 추가 분석 필요성이 증대하고 있으며, ③ 98년 혹은 09년 모집단 이후 변화된 모집단 특성을 반영하면서 광역시도 간 및 동읍면부로의 이동으로 인한 목표 표본 규모 기준의 표본 마모를 보완할 필요가 있으며, ④ 2015년에 통계청에서 처음 수행한 행정자료와 등록자료 기반의 등록센서스 모집단 기준의 패널 구축이 필요하기 때문이다. 그래서 17통합패널은 19차 패널과 변화된 모집단에서 새로운 패널을 추출해 추가함으로써 2015년 기준의 모집단에 대한 패널을 구축한다.

17통합패널의 목표 표본크기는 19차 조사의 응답 패널 7,000가구를 포함한 13,000가구이다. 따라서 17통합패널을 구축하기 위해 새롭게 추출하고자 하는 추가표본크기는 6,000가구이다. 17통합패널의 규모의 적절성을 살펴보기 위해 17통합패널 13,000가구를 대상으로 취업률, 이직률과 같은 고용 지표 추정을 위한 95% 신뢰수준에서의 최대 허용오차와 예상 상대표준오차(relative standard error)를 계산한 결과는 <표 2-44>와 같다. 여기서 상대표준오차는 $rse = \frac{se(\text{추정량})}{\text{추정량}} \times 100$으로 계산한다.

〈표 2-44〉 광역시도별, 동/읍면부별 및 주택유형별 예상 상대표준오차

		95% 신뢰수준 최대허용오차 (%)	예상 상대표준오차(RSE : p=0.5 기준, %)					
			목표 통합패널			현행 패널		
			계	아파트	일반	계	아파트	일반
전국	소계	0.86	0.88	1.22	1.26	1.19	1.66	1.72
	동부	0.95	0.97	1.41	1.34	1.35	1.96	1.87
	읍면부	2.00	2.05	2.44	3.77	2.56	3.16	4.36

첫째, 17통합패널 규모에 대한 취업률, 이직률과 같은 고용 지표 추정을 위한 95% 신뢰수준에서 최대 허용오차는 ∓0.8%p로 예상된다. 또한 동부와 읍면부로 보면, 동부가 1.0%p 이내, 읍면부가 2.0%p 이내로 예상된다.

둘째, 17통합패널 기준의 예상 상대표준오차는 1% 이내이다. 동읍면부별로는 동부가 1%, 읍면부가 2.1%이고, 주택유형별로는 아파트가 1.2%, 일반주택(보통)이 1.3% 이내로 예상된다. 특히, 읍면부의 아파트가 2.5%, 일반주택이 3.8% 이내로 예상되어 현행 패널에 비해 매우 안정적인 결과를 제공할 것으로 예상된다. 참고로, 현행 19차 기준의 노동패널(7,010가구 기준)에 대한 상대표준오차를 계산해 보면, 전국이 1.2%, 동부 1.4%, 읍면부가 2.6% 이내 수준이고, 조사구 특성별는 아파트 1.7%, 일반 1.7% 이내 수준이다. 또한 읍면부의 아파트가 3.2%, 일반이 4.4%로 동부보다 상대적으로 크지만 전국 단위에서는 5% 이내로 나타나 대체로 안정적인 변동을 나타내고 있다.

결론적으로 보면, 일반적인 전국 단위의 통계에 대한 상대표준오차가 5% 이내이면 매우 우수한 통계로 인식하므로 17통합패널의 규모는 매우 적절한 것으로 보인다.

2) 17통합패널의 표본 설계

17통합패널의 새로운 표본설계는 기본적으로 패널 구축의 표본설계 틀을 유지한다. 따라서 17통합패널의 표본설계에 대한 주된 초점은 첫째, 광역시도별 동읍면부 및 주택유형별 표본배분 방안이고, 둘째, 층별 추가 패널 추출을 위한 표본크기 결정과, 셋째, 추가 패널에 대한 추출 과정이다.

가) 표본추출방법

17통합패널의 표본설계에서 표본 추출 방법은 2단계 층화집락 확률비례계통추출법으로 표본을 추출한다. 17통합패널 설계의 층화 변수 및 기준은 현행 패널과 동일하게 적용하며 광역시도별 동읍면부 및 주택유형이다. 추가적으로 보통조사구 내에서 단독 주택과 연립/다세대 및 다가구의 단독주택 유형 및 주택면적(중형 기준)과 가구원수(1인 가구 비율)를 내재적 층화 변수로 사용한다. 내재적 층은 표본가구 선정과정에 반영하도록 한다.

〈표 2-45〉 17통합패널의 층화

	지역(1차)	2차 층화(동/읍면부)	3차 층화	층수
인구주택 총조사 병합조사구	7개 특별・광역시 1개 특별자치시	8개 특/광역/자치시별 동/읍면부(13개)	아파트(인천읍면부제외), 보통조사구(세종동부제외)	12 + 12 = 24
	9개 도	9개 도별 동/읍면부(18개)	아파트(제주읍면부제외), 보통조사구	17 + 18 = 35

17통합패널의 표본설계를 위한 집락은 평균 60가구의 병합조사구를 사용하고, 층수는 17개 광역시도, 2개 동/읍면부, 2개 조사구특성(주택유형)을 고려하고 일부 제외되는 층을 제외하면 모두 59개 층이다.

나) 층별 표본배분

17통합패널을 위한 표본설계에서는 광역시도별 동읍면부 및 주택유형별 표본 배분을 위해 비례배분, 제곱근비례배분, 선배정(광역시도별 100가구)+비례배분, 비례배분과 제곱근비례배분의 절충배분(절충1), 비례배분, 제곱근비례배분과 선배정+비례배분의 절충배분(절충2) 등을 검토한다.

<표 2-46>은 17통합패널의 표본 배분 방안별 배분 결과이다. 표본배분은 분석 목표, 통계 공표 단위의 오차 수준, 표본 관리 및 분석 용이성 등을 고려하여 결정하는 것이 바람직하다. 일반적으로 비례배분은 복합 조사 목적을 수행하거나 전국 단위 추정이 목적인 경우에 가장 널리 사용되는 배분 방법으로 지역별 표본의 구성비와 모집단의 구성비가 일치하도록 배분된다. 비례배분은 분석 및 추정과정에서 가중치의 변동이 크지 않

〈표 2-46〉 17통합패널의 배분 방안별 표본 규모

행정 구역	표본틀규모		표본배분 방안 검토					비례 (최종조정전)	
	가구수	%	비례	제곱근	선+비례	절충1	절충2	표본	구성비
전국	19,100,262	100.00	13,000	13,001	13,000	13,006	13,000	13,000	100.00
동부	15,597,553	81.66	10,618	9,319	9,612	9,972	9,849	10,610	81.62
읍면부	3,502,709	18.34	2,382	3,682	3,388	3,034	3,151	2,390	18.38
서울	3,874,286	20.28	2,637	1,568	2,392	2,103	2,199	2,640	20.31
동부	3,874,286	20.28	2,637	1,568	2,392	2,103	2,199	2,640	20.31
읍면부									
부산	1,331,861	6.97	906	920	888	913	904	900	6.92
동부	1,278,338	6.69	870	764	737	817	790	870	6.69
읍면부	53,523	0.28	36	156	151	96	114	30	0.23
대구	925,373	4.84	630	767	647	699	682	630	4.85
동부	858,596	4.50	585	600	506	593	564	590	4.54
읍면부	66,777	0.35	45	167	141	106	118	40	0.31
인천	1,042,844	5.46	710	814	717	762	747	710	5.46
동부	1,015,864	5.32	692	700	617	696	670	690	5.31
읍면부	26,980	0.14	18	114	100	66	77	20	0.15
광주	564,780	2.96	384	599	434	492	472	380	2.92
동부	564,780	2.96	384	599	434	492	472	380	2.92
읍면부									
대전	579,475	3.03	394	607	443	501	481	390	3.00
동부	579,475	3.03	394	607	443	501	481	390	3.00
읍면부									
울산	425,425	2.23	290	520	352	405	387	290	2.23
동부	345,116	1.81	235	351	237	293	274	240	1.85
읍면부	80,309	0.42	55	169	115	112	113	50	0.38
세종	67,129	0.35	46	206	140	126	131	50	0.38
동부	31,214	0.16	21	99	68	60	63	20	0.15
읍면부	35,915	0.19	25	107	72	66	68	30	0.23

〈표 2-46〉의 계속

행정 구역	표본틀규모		표본배분 방안 검토					비례 (최종조정전)	
	가구수	%	비례	제곱근	선+비례	절충1	절충2	표본	구성비
경기	4,462,411	23.36	3,037	1,683	2,740	2,361	2,487	3,030	23.31
동부	3,704,965	19.40	2,522	1,159	1,887	1,841	1,856	2,520	19.38
읍면부	757,446	3.97	515	524	853	520	631	510	3.92
강원	580,736	3.04	395	607	444	501	482	400	3.08
동부	352,444	1.85	240	336	246	288	274	240	1.85
읍면부	228,292	1.20	155	271	198	213	208	160	1.23
충북	591,441	3.10	403	613	450	508	488	400	3.08
동부	350,226	1.83	239	335	246	287	273	240	1.85
읍면부	241,215	1.26	164	278	204	221	215	160	1.23
충남	783,078	4.10	533	705	563	619	601	530	4.08
동부	346,747	1.82	236	332	265	284	278	230	1.77
읍면부	436,331	2.28	297	373	298	335	323	300	2.31
전북	700,118	3.67	477	667	514	573	552	480	3.69
동부	488,857	2.56	333	402	310	368	348	330	2.54
읍면부	211,261	1.11	144	265	204	205	204	150	1.15
전남	671,698	3.52	457	653	497	556	536	460	3.54
동부	291,610	1.53	198	305	232	252	245	200	1.54
읍면부	380,088	1.99	259	348	265	304	291	260	2.00
경북	1,039,384	5.44	707	812	715	760	745	710	5.46
동부	546,423	2.86	372	416	367	394	385	370	2.85
읍면부	492,961	2.58	335	396	348	366	360	340	2.62
경남	1,243,955	6.51	847	889	836	868	857	850	6.54
동부	808,753	4.23	551	513	482	532	515	550	4.23
읍면부	435,202	2.28	296	376	354	336	342	300	2.31
제주	216,268	1.13	147	371	228	259	249	150	1.15
동부	159,859	0.84	109	233	143	171	162	110	0.85
읍면부	56,409	0.30	38	138	85	88	87	40	0.31

으므로 추정 및 분석 과정에서 가중치의 영향이 작아지는 것이 장점이다. 반면, 지역별 혹은 층별 분석의 중요성이 보다 강조되는 조사라면 멱배분이 널리 사용된다. 멱배분은 층별 비교 및 분석을 위해 층별 오차 수준이 층간 비슷한 수준에서 유지되도록 표본을 배분하는 방법이다. 이 방법은 층간 표본크기 차이가 감소하도록 배분되는 것이 장점이지만 모집단의 구성비와는 일치하지 않는다는 점에 주의하고, 추정 및 분석 과정에서 가중치의 영향을 검토해야 한다. 멱배분 중 제곱근비례배분, 최소 표본크기 확보를 위한 선배분 후 나머지를 재배분하는 방법이 널리 사용된다. 그 외에도 다양한 목표를 달성하기 위해 여러 가지 배분방법으로 계산된 표본크기를 평균으로 절충한 절충배분도 사용하는데, 절충배분은 복합표본조사에서 전국 및 지역 추정의 목적을 동시에 달성하고자 할 때 사용 가능한 배분방법이다. 노동연구원 측과의 협의를 통해 중장기적인 종단연구를 위해 모집단 구성비와 동일하게 유지하고, 1차 추출단위인 조사구의 추출확률을 지역별로 비슷하게 유지하며, 또한 분석 및 추정 과정에서 가중치의 영향을 최소화하면서 지역별로 가중치의 영향을 비슷하게 부여하기 위해 비례배분의 방안으로 광역시도별 및 동읍면부별 표본을 배분하기로 결정하였다.

17통합패널의 목표 표본크기는 동부가 10,610가구, 읍면부가 2,390가구이다. 실제 층별 표본크기는 19차 현행 노동패널을 포함하는 수준에서 추가 추출 표본크기를 정한 후, 병합조사구당 10가구 추출을 위해 조사구를 결정하고, 표본 조사구당 10가구를 추출하여 층별 최종 표본크기를 결정한다.

다) 추가 표본크기 규모의 결정

17통합패널 구축을 위해 추가할 층별 표본크기는 <표 2-46>의 목표 표본크기에서 19차 현행 패널 크기를 제외한 표본크기를 2015년 추출틀에서 추가적으로 표본을 추가한다.

<표 2-47>은 광역시도별 동읍면부 및 주택유형별로 비례배분된 표본크기를 조정한 최종 표본크기이다. 병합조사구당 10가구를 추출하기 위해 조정하는 과정에서 동부는 10,610표본에서 10,600표본으로, 읍면부는

2,390표본에서 2,400표본으로 조정하였다. 19차 노동패널조사에 참여한 현재 패널을 제외한 과부족 표본을 추출해야 할 표본으로 조정하여 17통합패널 구축에 추가적으로 필요한 표본 규모를 결정한 것이다.

- 과부족 표본 = 목표 통합패널 규모(최종안) - 현재 운영 패널 규모 (19차)

<표 2-47>에서 보듯이 17통합패널 구축을 위해 추가적으로 추출해야 하는 표본은 5,988가구이다. <표 2-48>처럼 조사구당 10가구 추출을 위해 조정한다면 약 6,000가구이다. 구체적으로 추가할 표본가구수를 보면, 동부가 4,930가구, 읍면부가 1,070가구이다. 주택유형별로 보면, 보통조사구가 3,110가구이고, 아파트조사구가 2,890가구이다. 동부에서는 보통조사구가 2,330가구, 아파트조사구가 2,600가구이고, 읍면부는 보통조사구가 780가구, 아파트조사구가 290가구이다.

<표 2-48>은 17통합패널 구축을 위해 광역시도별 동읍면부 및 주택유형별로 병합조사구당 10가구 추출을 위해 조정한, 추가 추출하는 조사구수와 추가 가구수 현황이다. 기본적으로 2015년 인구주택총조사 추출틀에서 600개 조사구를 추출한다. 추출하는 병합조사구수는 동부가 493개, 읍면부가 107개이다. 아파트조사구가 289개이고, 보통조사구가 311개이다. 동부는 보통조사구가 233개, 아파트조사구가 260개이고, 읍면부는 보통조사구가 78개, 아파트조사구가 29개이다.

<표 2-49>는 17통합패널의 모집단, 통합패널 규모, 19차 노동패널 규모 및 추가 추출 규모를 요약한 표이다. 여기서 모집단은 2015년 인구주택총조사 조사구 추출틀의 모집단 가구수이다. 19차 노동패널의 규모 7,012가구, 추가 추출할 표본 규모가 6,000가구이며, 17통합패널의 규모는 13,012가구이다.

17통합패널의 표본 규모를 살펴보면, 동부 10,603가구, 읍면부 2,409가구이고, 보통조사구가 6,725가구, 아파트조사구가 6,287가구이다. 동부에서는 보통조사구가 5,014가구, 아파트조사구가 5,589가구이고, 읍면부에서는 보통조사구가 1,711가구, 아파트조사구가 698가구이다. 광역시도별로는 경기도가 3,034가구로 가장 많고, 서울이 2,647가구, 부산이 896가구,

〈표 2-47〉 통합패널 규모 및 과부족 표본 규모

행정 구역	목표 통합패널 규모(최종안)			현재 운영 패널 규모			과부족 표본 (추가 규모)		
	계	일반	아파트	계	일반	아파트	계	일반	아파트
전국	13,000	6,720	6,280	7,012	3,615	3,397	5,988	3,105	2,883
동부	10,600	5,020	5,580	5,673	2,684	2,989	4,927	2,336	2,591
읍면부	2,400	1,700	700	1,339	931	408	1,061	769	292
서울	2,640	1,540	1,100	1,297	744	553	1,343	796	547
동부	2,640	1,540	1,100	1,297	744	553	1,343	796	547
읍면부									
부산	900	430	470	526	276	250	374	154	220
동부	870	420	450	507	269	238	363	151	212
읍면부	30	10	20	19	7	12	11	3	8
대구	630	300	330	327	155	172	303	145	158
동부	580	280	300	303	139	164	277	141	136
읍면부	50	20	30	24	16	8	26	4	22
인천	710	330	380	413	189	224	297	141	156
동부	690	310	380	400	177	223	290	133	157
읍면부	20	20	0	13	12	1	7	8	-1
광주	380	140	240	186	70	116	194	70	124
동부	380	140	240	186	70	116	194	70	124
읍면부									
대전	390	180	210	186	89	97	204	91	113
동부	390	180	210	186	89	97	204	91	113
읍면부									
울산	290	120	170	190	79	111	100	41	59
동부	240	100	140	155	65	90	85	35	50
읍면부	50	20	30	35	14	21	15	6	9
세종	50	20	30	20	7	13	30	13	17
동부	20	0	20	11	0	11	9	0	9
읍면부	30	20	10	9	7	2	21	13	8

〈표 2-47〉의 계속

행정 구역	목표 통합패널 규모(최종안)			현재 운영 패널 규모			과부족 표본 (추가 규모)		
	계	일반	아파트	계	일반	아파트	계	일반	아파트
경기	3,030	1,380	1,650	1,534	637	897	1,496	743	753
동부	2,520	1,070	1,450	1,309	526	783	1,211	544	667
읍면부	510	310	200	225	111	114	285	199	86
강원	400	220	180	195	102	93	205	118	87
동부	240	100	140	130	46	84	110	54	56
읍면부	160	120	40	65	56	9	95	64	31
충북	400	210	190	247	139	108	153	71	82
동부	240	100	140	147	64	83	93	36	57
읍면부	160	110	50	100	75	25	60	35	25
충남	530	300	230	338	215	123	192	85	107
동부	230	90	140	145	72	73	85	18	67
읍면부	300	210	90	193	143	50	107	67	40
전북	480	260	220	265	161	104	215	99	116
동부	330	130	200	176	81	95	154	49	105
읍면부	150	130	20	89	80	9	61	50	11
전남	460	290	170	296	191	105	164	99	65
동부	200	80	120	127	50	77	73	30	43
읍면부	260	210	50	169	141	28	91	69	22
경북	710	440	270	408	246	162	302	194	108
동부	370	170	200	208	104	104	162	66	96
읍면부	340	270	70	200	142	58	140	128	12
경남	850	450	400	521	263	258	329	187	142
동부	550	240	310	338	151	187	212	89	123
읍면부	300	210	90	183	112	71	117	98	19
제주	150	110	40	63	52	11	87	58	29
동부	110	70	40	48	37	11	62	33	29
읍면부	40	40	0	15	15	0	25	25	0

〈표 2-48〉 통합패널 추가 조사구수 및 가구수 규모

행정구역	2017 추가 조사구수			2017 추가 가구수		
	계	일반	아파트	계	일반	아파트
전국	600	311	289	6,000	3,110	2,890
동부	493	233	260	4,930	2,330	2,600
읍면부	107	78	29	1,070	780	290
서울	135	80	55	1,350	800	550
동부	135	80	55	1,350	800	550
읍면부						
부산	37	15	22	370	150	220
동부	36	15	21	360	150	210
읍면부	1	0	1	10	0	10
대구	30	14	16	300	140	160
동부	28	14	14	280	140	140
읍면부	2	0	2	20	0	20
인천	30	14	16	300	140	160
동부	29	13	16	290	130	160
읍면부	1	1	0	10	10	0
광주	19	7	12	190	70	120
동부	19	7	12	190	70	120
읍면부						
대전	20	9	11	200	90	110
동부	20	9	11	200	90	110
읍면부						
울산	10	4	6	100	40	60
동부	8	3	5	80	30	50
읍면부	2	1	1	20	10	10
세종	3	1	2	30	10	20
동부	1	0	1	10	0	10
읍면부	2	1	1	20	10	10

〈표 2-48〉의 계속

행정구역	2017 추가 조사구수			2017 추가 가구수		
	계	일반	아파트	계	일반	아파트
경기	150	74	76	1,500	740	760
동부	121	54	67	1,210	540	670
읍면부	29	20	9	290	200	90
강원	20	11	9	200	110	90
동부	11	5	6	110	50	60
읍면부	9	6	3	90	60	30
충북	16	8	8	160	80	80
동부	10	4	6	100	40	60
읍면부	6	4	2	60	40	20
충남	19	9	10	190	90	100
동부	8	2	6	80	20	60
읍면부	11	7	4	110	70	40
전북	22	10	12	220	100	120
동부	16	5	11	160	50	110
읍면부	6	5	1	60	50	10
전남	16	10	6	160	100	60
동부	7	3	4	70	30	40
읍면부	9	7	2	90	70	20
경북	31	20	11	310	200	110
동부	17	7	10	170	70	100
읍면부	14	13	1	140	130	10
경남	33	19	14	330	190	140
동부	21	9	12	210	90	120
읍면부	12	10	2	120	100	20
제주	9	6	3	90	60	30
동부	6	3	3	60	30	30
읍면부	3	3	0	30	30	0

〈표 2-49〉 추출틀의 모집단, 통합패널규모, 현행 패널 규모 및 추가 추출 규모 (최종)

광역시도 및 동/읍면부		모집단			현행 패널 운영 규모			2017년 추가 추출 규모			통합패널 규모		
		가구수	일반	아파트	계	일반	아파트	계	일반	아파트	계	일반	아파트
전국	소계	19,100,262	9,866,097	9,234,165	7,012	3,615	3,397	6,000	3,110	2,890	13,012	6,725	6,287
	동부	15,597,553	7,407,000	8,190,553	5,673	2,684	2,989	4,930	2,330	2,600	10,603	5,014	5,589
	읍면부	3,502,709	2,459,097	1,043,612	1,339	931	408	1,070	780	290	2,409	1,711	698
서울	소계	3,874,286	2,262,416	1,611,870	1,297	744	553	1,350	800	550	2,647	1,544	1,103
	동부	3,874,286	2,262,416	1,611,870	1,297	744	553	1,350	800	550	2,647	1,544	1,103
	읍면부												
부산	소계	1,331,861	632,205	699,656	526	276	250	370	150	220	896	426	470
	동부	1,278,338	615,113	663,225	507	269	238	360	150	210	867	419	448
	읍면부	53,523	17,092	36,431	19	7	12	10	0	10	29	7	22
대구	소계	925,373	438,235	487,138	327	155	172	300	140	160	627	295	332
	동부	858,596	413,937	444,659	303	139	164	280	140	140	583	279	304
	읍면부	66,777	24,298	42,479	24	16	8	20	0	20	44	16	28
인천	소계	1,042,844	478,163	564,681	413	189	224	300	140	160	713	329	384
	동부	1,015,864	453,634	562,230	400	177	223	290	130	160	690	307	383
	읍면부	26,980	24,529	2,451	13	12	1	10	10	0	23	22	1
광주	소계	564,780	208,334	356,446	186	70	116	190	70	120	376	140	236
	동부	564,780	208,334	356,446	186	70	116	190	70	120	376	140	236
	읍면부												
대전	소계	579,475	261,789	317,686	186	89	97	200	90	110	386	179	207
	동부	579,475	261,789	317,686	186	89	97	200	90	110	386	179	207
	읍면부												
울산	소계	425,425	185,798	239,627	190	79	111	100	40	60	290	119	171
	동부	345,116	147,913	197,203	155	65	90	80	30	50	235	95	140
	읍면부	80,309	37,885	42,424	35	14	21	20	10	10	55	24	31
세종	소계	67,129	24,632	42,497	20	7	13	30	10	20	50	17	33
	동부	31,214	566	30,648	11	0	11	10	0	10	21	0	21
	읍면부	35,915	24,066	11,849	9	7	2	20	10	10	29	17	12

〈표 2-49〉의 계속

광역시도 및 동/읍면부		모집단			현행 패널 운영 규모			2017년 추가 추출 규모			통합패널 규모		
		가구수	일반	아파트	계	일반	아파트	계	일반	아파트	계	일반	아파트
경기	소계	4462,411	2,035,165	2,427,246	1,534	637	897	1,500	740	760	3,034	1,377	1,657
	동부	3704,965	1,575,146	2,129,819	1,309	526	783	1,210	540	670	2,519	1,066	1,453
	읍면부	757,446	460,019	297,427	225	111	114	290	200	90	515	311	204
강원	소계	580,736	325,035	255,701	195	102	93	200	110	90	395	212	183
	동부	352,444	149,941	202,503	130	46	84	110	50	60	240	96	144
	읍면부	228,292	175,094	53,198	65	56	9	90	60	30	155	116	39
충북	소계	591,441	319,922	271,519	247	139	108	160	80	80	407	219	188
	동부	350,226	148,967	201,259	147	64	83	100	40	60	247	104	143
	읍면부	241,215	170,955	70,260	100	75	25	60	40	20	160	115	45
충남	소계	783,078	447,530	335,548	338	215	123	190	90	100	528	305	223
	동부	346,747	141,473	205,274	145	72	73	80	20	60	225	92	133
	읍면부	436,331	306,057	130,274	193	143	50	110	70	40	303	213	90
전북	소계	700,118	376,505	323,613	265	161	104	220	100	120	485	261	224
	동부	488,857	196,185	292,672	176	81	95	160	50	110	336	131	205
	읍면부	211,261	180,320	30,941	89	80	9	60	50	10	149	130	19
전남	소계	671,698	412,676	259,022	296	191	105	160	100	60	456	291	165
	동부	291,610	110,109	181,501	127	50	77	70	30	40	197	80	117
	읍면부	380,088	302,567	77,521	169	141	28	90	70	20	259	211	48
경북	소계	1,039,384	641,509	397,875	408	246	162	310	200	110	718	446	272
	동부	546,423	257,312	289,111	208	104	104	170	70	100	378	174	204
	읍면부	492,961	384,197	108,764	200	142	58	140	130	10	340	272	68
경남	소계	1,243,955	656,919	587,036	521	263	258	330	190	140	851	453	398
	동부	808,753	358,869	449,884	338	151	187	210	90	120	548	241	307
	읍면부	435,202	298,050	137,152	183	112	71	120	100	20	303	212	91
제주	소계	216,268	159,264	57,004	63	52	11	90	60	30	153	112	41
	동부	159,859	105,296	54,563	48	37	11	60	30	30	108	67	41
	읍면부	56,409	53,968	2,441	15	15	0	30	30	0	45	45	0

〈표 2-50〉 통합패널의 최대허용오차 및 통합패널과 현행 패널의 예상 상대표준오차 비교

광역시도 및 동/읍면부		95% 신뢰수준 최대허용 오차 (%)	예상 상대표준오차(RSE : p=0.5 기준, %)					
			통합패널(최종 기준)			현행 패널		
			계	일반	아파트	계	일반	아파트
전국	소계	0.86	0.88	1.22	1.26	1.19	1.66	1.72
	동부	0.95	0.97	1.41	1.34	1.33	1.93	1.83
	읍면부	2.00	2.04	2.42	3.79	2.73	3.28	4.96
서울	소계	1.91	1.94	2.54	3.01	2.78	3.67	4.26
	동부	1.91	1.94	2.54	3.01	2.78	3.67	4.26
	읍면부							
부산	소계	3.27	3.34	4.85	4.62	4.36	6.03	6.34
	동부	3.32	3.40	4.89	4.73	4.44	6.11	6.49
	읍면부	18.19	18.89	40.82	21.82	23.57	40.82	30.15
대구	소계	3.91	4.00	5.83	5.49	5.54	8.06	7.65
	동부	4.04	4.14	6.00	5.74	5.75	8.51	7.83
	읍면부	15.69	15.24	25.81	19.24	20.85	25.81	37.79
인천	소계	3.68	3.75	5.52	5.11	4.93	7.29	6.70
	동부	3.73	3.81	5.71	5.11	5.01	7.54	6.71
	읍면부	22.47	21.31	21.81		28.86	30.14	
광주	소계	5.03	5.16	8.48	6.52	7.35	12.04	9.32
	동부	5.03	5.16	8.48	6.52	7.35	12.04	9.32
	읍면부							
대전	소계	4.97	5.09	7.49	6.97	7.35	10.66	10.20
	동부	4.97	5.09	7.49	6.97	7.35	10.66	10.20
	읍면부							
울산	소계	5.76	5.88	9.20	7.67	7.27	11.32	9.53
	동부	6.34	6.53	10.31	8.48	8.06	12.50	10.60
	읍면부	14.00	13.60	20.84	18.25	17.15	27.73	22.36
세종	소계	13.99	14.28	24.99	17.67	22.94	40.82	28.86
	동부	22.48	22.35		22.35			
	읍면부	18.19	18.89	24.99	30.14	35.35	40.82	99.99

〈표 2-50〉의 계속

광역시도 및 동/읍면부		95% 신뢰수준 최대허용 오차 (%)	예상 상대표준오차(RSE : p=0.5 기준, %)					
			통합패널(최종 기준)			현행 패널		
			계	일반	아파트	계	일반	아파트
경기	소계	1.78	1.82	2.69	2.46	2.55	3.96	3.34
	동부	1.95	1.99	3.06	2.62	2.76	4.36	3.58
	읍면부	4.34	4.41	5.68	7.02	6.68	9.53	9.41
강원	소계	4.90	5.04	6.88	7.41	7.18	9.95	10.42
	동부	6.34	6.47	10.26	8.36	8.80	14.90	10.97
	읍면부	7.77	8.06	9.32	16.22	12.50	13.48	35.35
충북	소계	4.90	4.96	6.77	7.31	6.37	8.51	9.67
	동부	6.34	6.37	9.85	8.39	8.27	12.60	11.04
	읍면부	7.77	7.93	9.36	15.07	10.05	11.62	20.41
충남	소계	4.26	4.35	5.73	6.71	5.45	6.83	9.05
	동부	6.47	6.68	10.48	8.70	8.33	11.86	11.78
	읍면부	5.67	5.75	6.87	10.60	7.22	8.39	14.28
전북	소계	4.48	4.54	6.20	6.69	6.15	7.90	9.85
	동부	5.40	5.46	8.77	7.00	7.56	11.18	10.31
	읍면부	8.03	8.22	8.80	23.56	10.66	11.25	35.35
전남	소계	4.57	4.69	5.87	7.81	5.82	7.25	9.80
	동부	6.94	7.14	11.25	9.28	8.91	14.28	11.47
	읍면부	6.09	6.22	6.90	14.58	7.71	8.45	19.24
경북	소계	3.68	3.73	4.74	6.07	4.96	6.39	7.88
	동부	5.10	5.15	7.60	7.02	6.95	9.85	9.85
	읍면부	5.32	5.43	6.07	12.21	7.09	8.42	13.24
경남	소계	3.36	3.43	4.70	5.02	4.38	6.18	6.24
	동부	4.18	4.27	6.45	5.71	5.45	8.16	7.33
	읍면부	5.67	5.75	6.88	10.54	7.41	9.49	11.95
제주	소계	8.03	8.11	9.49	15.81	12.70	14.00	31.62
	동부	9.38	9.66	12.31	15.81	14.58	16.66	31.62
	읍면부	15.69	15.07	15.07		26.72	26.72	

경남이 851가구의 순으로 많다, 반면 세종이 50가구로 가장 적고, 제주 153가구, 울산이 290가구의 순으로 적다.

라) 17통합패널의 타당성 검토

<표 2-49>에서 보듯이 추가 패널을 비례배분으로 배분하여 조사구당 10가구 추출을 하도록 조정하여 17통합패널의 광역시도별 동읍면부 및 주택유형별 표본크기를 배분해 결정하였다.

<표 2-50>은 층별 표본 규모에 대한 17통합패널의 타당성을 살펴보기 위해 ①95% 신뢰수준에서 모비율 추정 기준의 예상 최대허용오차(%)와 ②모비율 추정에서의 예상 상대표준오차(%)를 계산한 결과이다.

17통합패널의 95% 신뢰수준에서 모비율 추정의 최대허용오차는 ∓0.86%p로 예상된다. 동부가 ∓0.95%p, 읍면부가 ∓2.00%p로 예상되어 매우 작은 오차의 결과가 예상된다. 광역지역별로 보면, 경기가 ∓1.78%p로 가장 작으며, 세종이 ∓13.99%p로 가장 크다. 세종은 모집단 규모가 작아 14%p 수준이지만 세종시를 제외한 16개 광역시도는 모두 ∓8.0%p 이내의 결과가 예상된다.

17통합패널에 대한 모비율 추정의 상대표준오차는 0.88%로 예상된다. 동읍면부별로는 동부가 0.97%, 읍면부가 2.04%이고, 주택유형별로는 보통조사구가 1.22%, 아파트조사구가 1.26%로 예상된다. 동부의 경우, 보통조사구가 1.41%, 아파트조사구가 1.34%이고, 읍면부에서는 보통조사구가 2.42%, 아파트조사구가 3.79%이다. 광역시도별로는 경기도가 1.82%로 가장 작으며, 세종시가 14.28%로 가장 크다. 세종시를 제외한 16개 광역시도는 모두 8.2% 이내의 결과가 예상된다. 참고로, 19차 노동패널(현행 패널)의 상대표준오차는 1.19%이고, 동읍면부별로는 동부가 1.33%, 읍면부가 2.73%이고, 주택유형별로는 보통조사구가 1.66%, 아파트조사구가 1.72%이다.

17통합패널에 대한 광역시도별 동읍면부 및 주택유형별 표본크기에 대한 타당성을 모비율 추정의 최대허용오차 및 상대표준오차 관점에서 살펴보면, 광역시도별에서는 세종시를 제외한 16개 광역시도가 8.2% 이내에서 안정적인 결과를 제공하고 있으며, 동읍면부별 및 주택유형별에서

는 4% 이내로 예상되어 현행 19차 노동패널보다 매우 안정적인 결과를 제공할 것으로 예상된다.

마) 표본추출과정

앞서 설명한 바와 같이 17통합패널 구축을 위한 추가 패널의 표본 추출 방법은 2단계 층화집락 확률비례계통추출법으로 표본을 추출한다.

2단계 층화집락 확률비례계통추출법으로 추출하기 위한 조사구의 정렬 순서는 층-주거면적(중형 이상)-1인 가구수 비율-시군구-(단독주택 비율)-가구주연령대-자가비율의 순으로 정렬한다.

추가 패널의 표본추출 과정은 다음과 같다.

1단계는 층별 병합조사구를 가구수에 비례하는 pps 추출법으로 표본조사구를 추출하고,

2단계는 표본조사구 내에서 10가구를 계통 추출하여 표본가구를 추출한다.

바) 패널 선정 및 예비표본 추출

17통합패널 구축을 위해 추가해야 할 패널 추출을 위한 표본설계에서 표본 가구 관리를 위해 다음과 같은 사항을 패널 선정 과정에서 고려할 것을 검토하기로 한다.

첫째, 가구원수 관리에 대한 사항으로 2015년 인구주택총조사에서 나타난 1인 가구 특성을 반영하기 위해 동부에서 젊은층의 1인 가구 특성이 반영되도록 한다. 또한 19차 패널에서 다수를 차지하고 있는 4인 가족 비율(19차 기준 24.7%, 모집단 18.8%)을 통제하여 패널 가구를 선정하도록 한다.

둘째, 표본조사구의 대체를 위해 예비 표본조사구를 3배수 추출하기로 한다. 그러나 기본적으로 조사구당 10가구를 추출하여 모두 패널로 영입할 것을 목표로 하는 것이 기본 원칙이고, 조사구 내에서 7가구 이상 패널 영입이 가능하다면 표본조사구를 대체하지 않도록 한다. 이를 위해 조사구의 가구수가 많은 조사구에서는 13가구까지 추출해 패널로 영입하는 방안의 검토가 필요하다. 결과적으로 표본조사구의 대체는 재개발 및 재건축으로 인한 조사구 손실, 접촉이 불가능한 조사구 등 표본가구 접촉이

불가능한 상황을 제외하고는 표본조사구 내에서 가구를 대체하도록 하고, 표본조사구의 대체를 최소화하도록 한다[2].

셋째, 예비표본 조사구의 추출은 원표본 조사구와 동일한 지역에서 추출하도록 한다. 하지만 표본틀에서 동읍면코드를 제공하지 않으므로 동일 시군구 내에서 추출하여 예비조사구를 제공하기로 한다.

5. 가중치와 추정식

4항에서 17통합패널의 광역시도별 동읍면부 및 주택유형별 목표 표본크기를 결정하고, 19차 현행 노동패널의 현황에 근거해 17통합패널을 위해 추가해야 할 패널 추출을 위한 표본설계 방안을 살펴보았다.

본항에서는 17통합패널의 가중치 산출방안과 추정식을 설명하고자 한다. 한국노동패널조사의 가중치는 종단면 개인 가중치, 횡단면 가구 가중치와 개인 가중치를 산출해 제공한다. 일반적으로 17통합패널과 같이 기존 패널을 확대해 추가할 때 가중치 산출 방안은 다음과 같은 두 가지의 산출 방안이 가능하다.

첫째, 17년 조사에 참여한 모든 응답 가구 및 가구원만을 17통합패널의 원패널로 정의해 기존의 각 패널의 추출 과정을 무시하고 원패널 중 응답 완료된 패널에 대해서만 사후층화 가중치로 17년 횡단가중치를 산출하는 방안(방안 1),

둘째, 17통합패널은 1998년 원패널, 2009년 추가 패널 및 2017년 추가 패널로 구성되므로 패널 추출 과정으로 각각 고려한 패널가중치를 별도로 산출하여 통합패널을 위한 통합가중치를 산출하는 방안(방안 2).

방안 1은 가중치 산출 과정이 용이하지만 패널의 추출 과정을 고려하지 않는 점이 단점이 되고, 방안 2는 각 패널의 추출 특성을 고려하지만 특성이 동일한 층의 가구 및 가구원 가중치가 98패널, 09년 추가 패널 및

2) 실제 시범조사의 조사지침에서는 조사구의 가구수가 30가구 이상인 경우는 최대 14가구, 최소 6가구를 진행하면 성공한 조사구로 인정하였다. 조사구 가구수가 30가구 미만인 경우에는 최대 10가구, 최소 4가구를 진행하면 성공한 것으로 간주하였다.

17년 추가 패널마다 다른 가중치를 가지므로 이용자에게 혼란이 야기되는 점이 단점이다. 특히 17년 추가 패널의 통합적인 가구 특성을 고려한다면 가중치 산출이 더 어려워지는 단점도 나타난다. 그러므로 방안 1에서는 조사완료 후 2015년 및 2017년 기준의 모집단으로 사후층화가중치만을 부여하면 되고, 방안 2는 09통합패널은 17년 횡단가중치를 사용하고, 17년 추가 패널에 대해서만 2015년 모집단 기준의 가중치를 산출하고, 패널 특성을 고려하여 통합가중치를 산출하면 된다.

17통합패널의 가중치는 기본적으로 현행 노동패널의 가중치와 동일한 과정으로 산출하도록 한다. 한국노동패널조사는 98년 최초 수행 이후 2009년도에 09통합패널로 1차 확대하였고, 2차로 17통합패널을 구축하는 상황이다. 따라서 17통합패널 가중치 산출은 09통합패널과 동일하게 적용하기로 한다. 17통합패널 가중치를 산출하기 위해 다음과 같이 몇 가지 사항을 가정한다.

첫째, 2017년 추가 패널을 포함하여 기존 09통합패널의 20차 조사(2017년)에 참여해 2017년에 조사에 응답한 모든 응답 가구원 및 가구를 17통합패널의 원패널 가구로 정의한다.

둘째, 17년 추가 패널은 독자 패널로 관리되는 것이 아니라, 2009년 이후 모집단이 변화되고, 또한 기존 09통합패널의 마모, 탈락, 지역 간 이주 등 기존 패널의 변화로 인해 통계분석 단위 간 표본 확보 및 추가 요구사항에 부합하기 위해 패널을 추가하는 상황으로, 17년 추가 패널을 포함한 17통합패널로 운영하므로 2017년이 17통합패널의 원패널로 선정한다. 즉 17년 추가 패널을 추가함으로써 노동패널은 98패널, 09년에 전국으로 확대하기 위해 추가한 09통합패널, 17년 모집단 변화와 표본 마모 및 변화를 보완하기 위해 추가한 17통합패널로 구성하게 된다.

셋째, 노동패널 가구 및 가구원 가중치는 종단 및 횡단 가중치를 부여하는데 기존 98패널 및 09통합패널에 대한 가중치는 한국노동연구원에서 현재 가중치를 산출하고 있으므로 본 연구에서는 고려하지 않는다.

넷째, 노동패널에서 가구원 가중치는 가구원에게 가구 가중치를 동일하게 부여해 가구원 가중치를 산출하여 무응답 및 사후층화 조정 가중계수로 조정하므로 본 연구에서는 가구 가중치 산출 방안만을 제시하기로

한다.

다섯째, 17통합패널의 원년(1차) 패널 가중치는 종단가중치와 횡단가중치가 동일하므로 횡단가중치 산출 방법만을 제시하기로 한다.

따라서 본 연구에서는 2017년 조사에 참여한 가구 및 가구원만을 원패널로 선정하여 17통합패널에 대한 가중치 산출 방안만을 고려하기로 한다.

가. 가중치

1) 가중치 산출 개요

대부분의 표본조사가 확률적 표본추출 방법론으로 설계함에도 불구하고 추출된 표본 단위의 추출 과정 및 응답 여부에 따라 응답 표본 단위간 불균등이 초래된다. 따라서 모집단 변화 및 응답 표본 분포와 모집단 분포의 차이를 보정하기 위해 가중치를 부여해 추정 및 분석하는 것이 편향을 제거하고 정확성을 제고할 수 있으므로 바람직하다.

일반적으로 가중치는 표본 자료를 모집단 관점에서 분석하도록 확대시켜 주는 확장계수이고, 모집단 대표성을 보여주는 수치이다. 이론적으로 표본 추출 및 자료수집과정, 모집단의 커버리지 등 조사된 결과를 표본추출에 따른 추출률의 차이와 응답률 및 모집단에 대한 정보 등을 이용하여 가중치를 부여하면 모집단의 구조와 표본 구조는 일치하게 되고, 편향을 제거함으로써 추정의 정확도를 높일 수 있다. 그러나 표본단위들의 가중치는 추정량의 분산 추정에 영향을 주며, 가중치의 변동이 심하면 분산이 증가하는 단점이 나타남에 주의해야 한다.

추정단계에서 가중치를 이용하면 모집단에 대한 특성치인 모수에 대한 비편향 추정량(unbiased estimator)을 얻게 되지만, 분석 과정에서 가중치를 무시하고 분석한 추정치는 심각한 편향(bias)이 발생하게 된다. 표본크기가 큰 대규모 표본조사에서 모수를 추정하는 것이 목표라면 비편향 추정량을 얻기 위해 표본설계 과정 및 가중치를 반드시 부여하여 추정해야 한다.

2) 가구 가중치 산출 과정

일반적으로 가중치를 산출하기 위해서는 모집단 정보, 표본추출과정 및 추출 단위, 추출률과 응답률 등의 조사 관리자료 정보가 필요하다.

17통합패널 구축을 위해 추가해야 할 패널 추출을 위한 설계시점의 모집단은 2015년 11월 기준의 「인구주택총조사」이고, 조사 및 분석을 위한 사후층화 조정을 위한 조사시점의 모집단은 2017년 「추계가구」이다.

17통합패널 구축을 위한 추가 패널의 표본추출은 2단 층화집락 확률비례 계통추출법이고, 1차 추출단위는 조사구로 조사구의 가구수 크기에 비례하는 확률비례 계통추출법으로 추출하고, 2차 추출단위는 가구로 계통추출로 추출한다.

일반적으로 표본이론에서 표본조사의 가중치는 ① 설계가중치, ② 무응답에 대한 보정, ③ 모집단 정보를 이용한 사후층화 조정 등의 과정을 통해 최종 가중치를 산출한다.

설계가중치는 표본추출을 위해 2단 층화집락 확률비례 계통추출법을 적용함에 따라 나타나는 층별 조사구 내의 가구별 추출률의 차이를 반영하기 위해 부여하는 가중치이다. 설계가중치는 표본추출과정을 표현하는 가중치로 최종 표본단위의 추출 단계별 추출률을 곱한 표본단위의 표본추출률의 역수로 계산하는 가중치이다. 17통합패널에서는 최종적으로 가구가 표본단위로 추출되는데, 가구 가중치는 층별로 조사구 추출 및 조사구 내 가구 추출을 곱해 산출한다.

무응답 보정 가중치는 표본가구의 부재, 조사 참여 거부 등으로 인해 응답하지 않아서 발생되는 정보의 손실과 추정의 편향을 응답가구로 하여금 보정하도록 부여하는 가중치이다. 기본적으로 무응답 보정 가중치는 표본단위의 응답률의 역수로 산출한다. 표본단위의 응답률은 ① 무응답 조정 층별 응답률, ② 응답여부에 영향을 미치는 주요한 가구 특성변수 등을 이용한 로지스틱회귀모형을 통해 추정한 개별 표본의 추정 응답확률로 추정한다. 이론적으로 응답층과 무응답층의 특성 차이가 크지 않거나 응답률이 높으면 무응답 보정 가중치를 부여함으로써 무응답으로 인한 편향을 감소시킬 수 있다. 노동패널에서는 종단가중치 산출 시 전년도 표본이 조사연도에 응답하였는지 여부로 무응답 조정 가중치를 산출

한다. 17통합패널에서는 원가구원의 무응답에 대해 무응답 층별 응답률로 무응답 보정 가중치를 계산한다.

사후층화 가중치는 추정 모집단 분포와 실제 모집단 분포를 일치하도록 조정하는 가중치를 말한다. 사후층화 조정 가중치는 설계가중치와 무응답 보정 가중치를 부여해 계산한 가중치의 합은 모집단의 크기와 일치해야 하는데, 설계시점의 표본틀에서 누락된 모집단 특성을 반영하거나 모집단이 변화하여 설계시점의 모집단과 조사 및 분석 시점의 모집단 특성이나 분포에 차이가 있을 때 모집단의 커버리지 및 모집단 변화를 조정하기 위해 부여하는 가중치이다. 사후층화 조정 가중치는 ① 모집단의 정보가 완전히 알려져 있다면 사후 층별 추정 모집단 분포와 실제 모집단 분포를 일치하도록 조정하는 일반적인 사후층화 조정 가중치(Post-Stratification Adjustment : PSA), ② 주요 변수의 주변 합(marginal total)과 같이 모집단의 정보가 부분적으로 알려져 있다면 알려진 주변 합 정보를 이용해 반복 과정으로 층별 결합확률이나 모집단 크기를 추정하여 추정 모집단을 설정한 후 이를 기준으로 조정하는 갈퀴비 조정 가중치(raking ratio weight : 일부 RIM 가중이라고도 함), ③ 모집단에 대한 정보가 거의 존재하지 않거나 기준 모집단 정보로 사용하기에 정확성이 낮은 경우라면 다른 대체적인 정보를 이용하여 추정 모집단을 설정한 후 이를 기준으로 조정 가중치를 산출하는 벤치마킹 가중치(Bench-marking weight)를 이용하여 산출한다. 이론적으로 사후층화 조정 가중치를 부여하면 모집단 및 표본틀의 커버리지로 인한 편향 및 모집단 변화에 따른 편향뿐만 아니라 추정량의 분산도 감소하게 되므로 표본조사에서는 사후층화 조정 가중치를 부여하는 것이 바람직하다. 17통합패널에서 표본 설계 시점은 2015년 인구주택총조사 자료를 사용하지만 자료수집 및 분석 기준 시점은 2017년 추계 가구 모집단을 기준으로 하는 지역(광역시도, 동읍면부), 가구원수, 주택유형 및 난방 방식을 갈퀴비 조정을 위한 변수로 선정하여 갈퀴비 조정 가중치로 사후층화 조정 가중치를 계산한다.

가중치 작성 과정을 설명하기 위한 기호들을 정리하면 다음과 같다.

- L : 층의 수

- N_h : 층 h의 모집단조사구 수
- n_h : 층 h의 표본조사구 수
- S_{hi} : 층 h의 i번째 조사구에 대한 크기 측도(해당 조사구의 총 가구수)
- $S_h = \sum_{i=1}^{N_h} S_{hi}$: 층 h에서 크기의 측도에 대한 총합
- M_{hi} : 층 h의 i번째 조사구 내 가구수
- m_{hi} : 층 h의 i번째 조사구 내 표본가구수(응답+거절)
- r_{hi} : 층 h의 i번째 표본조사구 내 조사완료 가구수(응답)

가) 설계가중치

설계가중치는 2단계 층화집락 확률비례 계통추출법을 반영하기 위해 1단계에서 확률비례 추출한 표본조사구에 대한 표본추출률의 역수와 2단계에서 표본조사구 내에서 가구조사의 접촉(추출)률의 역수를 곱하여 다음과 같이 산출한다.

- 설계가중치 $= \frac{S_h}{n_h\ S_{hi}} \times \frac{M_{hi}}{m_{hi}}$

원칙적으로 각 표본조사구에서는 평균 10가구씩을 표본으로 조사하였기 때문에 $m_{hi} = 10$이고, $S_{hi} \approx M_{hi}$를 가정할 수 있는 경우(표본추출틀 상의 조사구 내 가구수와 실제 가구수에 차이가 작은 경우), 각 층에서 설계가중치는 해당 층 내에서 일정한 값이 되어 설계가중치 $= \frac{S_h}{n_h \times 10}$으로 표현할 수 있다. 이때, 각 지역 내 층에서 표본가구들은 모두 동일한 자체 가중 설계가중치를 갖게 된다(epsem).

실제 조사과정에서 표본 추출틀의 조사구 내 가구수와 실제 가구수에 차이가 있어 각 층 내에서 설계가중치는 동일하지는 않지만 조사구별 가구수가 평균 60가구로 일정하므로 조사구 내 가구 가중치는 비슷한 값을 갖는다.

나) 무응답 보정

기본적으로 층(광역시도×동읍면부, 주택유형) 내 표본가구의 설계가

중치는 원칙적으로 동일하다. 17통합패널의 무응답 조정은 무응답 조정 층별 응답률로 산출하기 때문에 무응답 조정 층별 무응답 조정 가중치는 다음과 같이 계산한다. 여기서 무응답 층은 설계 층과 동일하게 사용해도 되지만 응답률에 영향을 주는 변수를 이용한 층으로 설정하는 것이 바람직하다.

$$-\text{무응답 조정계수} = \frac{m_{hi}}{r_{hi}}$$

다) 2017년 「추계 가구」를 이용한 사후층화 조정

모집단 커버리지 편향 및 모집단 변화를 위한 사후층화 조정 가중치는 최신 모집단 정보를 이용하는 것이 모집단 커버리지 편향 제거 및 추정량의 분산 감소 등 추정의 정확도를 높이게 된다. 17통합패널의 설계 기준 모집단은 2015년 「인구주택총조사」이고, 사후층화 조정 가중치를 위한 기준 모집단은 「2017년 추계 가구」이다. 추정 및 분석을 위해 최신 모집단인 2017년 추계 가구 자료를 이용해야 하는데, 추계 가구는 가구수를 광역시도, 동읍면부별 총계만 공표하는 부분적인 모집단 정보를 제공하게 되므로 2017년 현재 추계가구분포에 대해 갈퀴비 조정(Raking Ratio Adjustment Method)을 적용하여 가중치를 계산한다.

갈퀴비 조정(Raking Ratio Adjustment Method)에 활용한 변수는 추계가구분포에 대해서 (17개시도 * 동・읍/면별(2))변수들로 선정한다.

라) 최종 가중치

17통합패널의 가구 최종 (횡단)가중치는 다음과 같이 산정한다.

최종 가중치=설계가중치×무응답 조정 가중치×사후층화 조정 가중치

마) 극단 가중치 조정

일반적으로 가중치는 편향을 제거하지만 가중치의 과도한 변동은 추정량의 분산을 크게 해 추정의 정확도를 떨어뜨리기 때문에, 지나치게 크거나 작은 몇 개의 가중치가 존재한다면 매우 잘 설계되어 수행된 통계조사에서 조사의 정확성을 낮추게 되는 현상이 발생하기도 한다.

극단 가중치는 지나치게 크거나 작은 가중치를 의미하며, 극단 가중치는 가중치의 분포를 통해 이상치 여부를 판단하는 것이 바람직하다.

- 가중치가 정규분포를 따른다면, ① $\overline{w} \mp 2 \times s_w$ 혹은 $\overline{w} \mp 3 \times s_w$, ② 상위 5%(혹은 3%)의 경계를 초과하는 가중치를 극단가중치로 판단할 수 있다.
- 가중치가 정규분포를 확신할 수 없다면, ① $\tilde{w} \mp 1.5 \times IQR$ 혹은 $\tilde{w} \mp 3 \times IQR$, ② 최대값과 최소값의 비가 5배(혹은 10)배), ③ 평균의 2~3배의 경계를 초과하는 가중치를 극단 가중치로 판단할 수 있다.

기본적으로 극단 가중치는 표본크기 및 배분 과정에서 규모가 큰 모집단에 비해 표본수가 지나치게 적을 때 발생하는 것이므로 설계 단계에서 통제가 가능하지만, 실사 가정의 무응답 및 모집단 분포의 변화로 발생하기도 한다. 극단 가중치는 설계가중치, 무응답 조정, 사후층화 조정 등 가중치 작성 단계별로 조정하는 것이 바람직하며, 극단 가중치는 조사 완료 후 조정하는 방법이 사용되고 있는데, 대표적인 방법이 윈저화(winsorization : 극단 가중치를 가중치 상한 내에서 가장 유사한 가중치로 재부여하는 방법)이다.

극단 가중치가 존재하는 경우, 극단 가중치를 절단함으로써 손실되는 정보의 양을 조정하는 가중치를 계산해 최종 가중치를 조정한다. 이때, 극단 가중치를 절단해 조정한 가중치는 전체 추정에만 사용하고, 기본적인 추정은 극단 가중치 조정 전 최종 가중치를 사용해 추정하고, 또한 극단 가중치를 조정함으로써 얻게 되는 가중치 조정 효과를 보고서에 표현하는 것이 바람직하다.

$$\mathrm{deff}_w = 1 + cv_w^2$$

17통합패널에서의 가중치 조정은 winsorization 방법을 사용하고, 전체 응답자의 가중치 분포를 고려하여 $\overline{w} \mp 2 \times s_w$ 를 기준으로 극단 가중치 경계를 설정하는 방안을 권유한다(이때, 가중치가 분산에 미치는 영향을 검토하여 결정하는 방안도 고려할 필요가 있다).

나. 추정식

2017통합패널을 위한 새로운 표본설계에서 추정식은 17통합패널의 최종 가중치를 이용한 추정식을 제안한다.

• 총합 추정치

층별 특성 X의 합계추정치 $\hat{X}_{gh}$, 시도별 합계추정치 $\hat{X}_g$ 및 전국추정치 $\hat{X}$라고 하면,

$$\hat{X}_{gh} = \sum_i X_{ghi} = \sum_i \sum_j w_{ghij} x_{gij}$$

$$\hat{X}_g = \sum_h \hat{X}_{gh}$$

$$\hat{X} = \sum_g X_g$$

여기서

$\hat{X}$: 특성 X에 대한 값의 합계추정치
x : 특성 X에 대하여 조사한 값
w : 가중치(설계가중치, 무응답 조정 및 사후층화 가중치 반영)
g : 시도를 나타내는 첨자
h : 층(동읍면부, 조사구특성/주택유형)을 나타내는 첨자
i : 표본조사구를 나타내는 첨자
j : 조사된 가구를 나타내는 첨자

층별 추정치에 대한 분산 $Var(\hat{X}_{gh})$, 표준오차 $Se(\widehat{X_{gh}})$ 및 상대표준오차 $RSE(\widehat{X_{gh}})$ 추정치는 다음과 같다.

$$Var(\widehat{X_{gh}}) = \frac{n_{gh}}{n_{gh}-1} \sum_i \left(x_{ghi} - \bar{x}_{gh}\right)^2$$

여기서 $x_{ghi} = \sum_j w_{ghij} x_{ghij}$, $\bar{x}_{gh} = \sum_i x_{ghi} / n_{gh}$

$$Se(\widehat{X_{gh}}) = [Var(\widehat{X_{gh}})]^{1/2}$$

$$RSE(\widehat{X_{gh}}) = [Se(\widehat{X_{gh}}) / \widehat{X_{gh}}] \times 100$$

그리고 광역시도별 추정치 $\hat{X}_g$에 대한 분산 $Var(\hat{X}_g)$, 표준오차 $Se(\hat{X}_g)$

및 상대표준오차 $RSE(\widehat{X_g})$ 추정치는 다음과 같다.

$$Var(\widehat{X_g}) = \sum_h Var(\widehat{X_{gh}})$$
$$Se(\hat{X}_g) = [Var(\hat{X}_g)]^{1/2}$$
$$RSE(\widehat{X_g}) = [Se(\widehat{X_g})/\widehat{X_g}] \times 100$$

최종적으로 전국 추정치 $\hat{X}$에 대한 분산 $Var(\hat{X})$, 표준오차 $Se(\hat{X})$ 및 상대표준오차 $RSE(\hat{X})$ 추정치는 다음과 같이 추정한다.

$$Var(\hat{X}) = \sum_g Var(\widehat{X_g})$$
$$Se(\hat{X}) = [Var(\hat{X})]^{1/2}$$
$$RSE(\hat{X}) = [Se(\hat{X})/\hat{X}] \times 100$$

• 전국 기준 평균 추정치 및 분산 추정치

17통합패널에서 주요 추정 모수가 모평균인 경우는 표본단위의 가중치를 사용한 가중 평균으로 계산되며, 가중치를 이용한 모평균 추정량은 다음과 같다.

$$\hat{\bar{X}} = \frac{\sum_{g=1}^{G}\sum_{h=1}^{L}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} w_{ghij} x_{ghij}}{\sum_{g=1}^{G}\sum_{h=1}^{L}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} w_{ghij}} = \frac{\sum_{g=1}^{G}\sum_{h=1}^{L}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} w_{ghij} x_{ghij}}{w_{....}}$$

여기서 모비율은 측정값이 1 또는 0의 값을 갖는 점이 모평균과 다르지만 모비율은 모평균과 동일한 수식으로 표현한다.

- 모비율 $x_{ghij} = \begin{cases} 1, \text{조건 만족} \\ 0, o.w. \end{cases}$

그리고 평균(비율) 추정량의 분산 추정량은 표본 추출과정에서 이용한 층화 및 집락 등의 표본설계를 반영하여 추정하는데, 모평균 및 모비율 추정량의 분산 추정량은 다음과 같이 계산한다. 참고로, 모평균과 모비율 추정량의 분산 추정량도 추정량과 마찬가지로 동일하게 추정하면 된다.

$$Var(\hat{\overline{X}}) = \sum_{g=1}^{G}\sum_{h=1}^{L}\frac{n_{gh}(1-f_{gh})}{n_{gh}-1}\sum_{i=1}^{n_h}(e_{ghi.} - \overline{e_{gh..}})^2$$

여기서 $f_{gh} = n_{gh}/N_{gh}$, $e_{ghi.} = \left(\sum_{j=1}^{m_{hi}} w_{ghij}(x_{ghij} - \overline{x}_g)\right)/w_{g...}$,

$$\overline{e_{gh..}} = \left(\sum_{i=1}^{n_{gh}} e_{ghi.}\right)/n_{gh}$$

여기서 w_{ghij}는 각 응답자에 부여된 가중치이고, y_{ghij}는 표본단위의 응답자료이다. 그리고 L은 층의 수, n_{gh}는 광역시도 g의 층 h에서 1차 표본추출단위인 표본조사구의 수, n_h는 층 h에서의 1차 추출단위(PSU)인 조사구수, m_{hi}는 층 h내의 i번째 표본조사구의 응답 가구수, $w_{...} = \sum_{h=1}^{L}\sum_{i=1}^{n_h}\sum_{j=1}^{m_{hi}} w_{hij}$은 전체 응답 가구에 대한 가중치의 합계(=모집단 크기)이다.

또한, 모평균 및 모비율 추정에 대한 표준오차(standard error), 상대표준오차(relative standard error), 95% 신뢰수준 오차의 한계는 다음과 같다.

- $SE(\overline{x}) = \sqrt{Var(\hat{\overline{X}})}$,
- $RSE(\overline{x}) = \frac{SE(\overline{x})}{\overline{x}} \times 100(\%)$,
- 오차의 한계 $= \mp 1.96 \times \sqrt{Var(\hat{\overline{X}})} = \mp 1.96 \times SE(\overline{x})$

17통합패널에서 가중치를 반영한 추정을 위해서는 복합표본조사자료를 분석하는 전문 패키지를 사용하는 것이 필요하다. 17통합패널의 추정 및 분석을 위해 SPSS의 복합표본(Complex Samples), SAS의 survey 프로시저(proc surveymeans 등), Stata, R의 survey 패키지 등의 사용을 권유한다.

〈부표 1〉 지역별, 성별 및 연령에 대한 추계자료와 패널자료의 비율 차이

〈전 국〉

(단위 : %)

연령	전체	성별	
		남자	여자
15~19세	1.07	0.85	1.23
20~24세	3.15	3.93	2.39
25~29세	2.24	2.48	2.00
30~34세	1.26	1.26	1.22
35~39세	-1.49	-1.48	-1.53
40~44세	-0.65	-1.23	-0.12
45~49세	0.77	0.11	1.37
50~54세	1.24	1.27	1.18
55~59세	0.07	0.17	-0.02
60~64세	-1.07	-1.03	-1.11
65~69세	-1.92	-1.76	-2.06
70~74세	-1.37	-1.82	-0.94
75~79세	-1.91	-1.55	-2.22
80~84세	-0.86	-0.87	-0.81
85~89세	-0.38	-0.24	-0.48
90~94세	-0.12	-0.1	-0.12
95~99세	0.00	-0.01	0.02
100세 이상	0.00	0.00	-0.02

〈서 울〉

(단위 : %)

연령	전체	성별	
		남자	여자
15~19세	0.70	0.40	0.94
20~24세	3.46	3.29	3.61
25~29세	3.47	3.48	3.46
30~34세	2.32	2.91	1.78
35~39세	0.07	-0.34	0.43
40~44세	-1.27	-1.79	-0.83
45~49세	-0.28	-1.00	0.36
50~54세	1.21	1.89	0.60
55~59세	-0.54	-0.36	-0.71
60~64세	-2.29	-1.60	-2.91
65~69세	-1.92	-2.20	-1.65
70~74세	-1.18	-0.87	-1.45
75~79세	-2.42	-2.46	-2.36
80~84세	-0.86	-0.73	-0.97
85~89세	-0.50	-0.57	-0.42
90~94세	-0.02	-0.06	0.04
95~99세	0.05	0.02	0.07
100세 이상	0.01	0.00	0.01

〈부 산〉

(단위 : %)

연령	전체	성별	
		남자	여자
15~19세	1.72	2.14	1.33
20~24세	2.44	3.56	1.64
25~29세	1.00	1.43	0.32
30~34세	-1.65	-1.90	-0.95
35~39세	-0.24	-1.32	-0.11
40~44세	0.95	0.19	1.14
45~49세	1.79	1.32	1.87
50~54세	2.78	2.08	3.24
55~59세	-1.45	0.36	-2.86
60~64세	-1.22	-2.77	0.28
65~69세	-1.85	-1.92	-1.22
70~74세	-0.77	-1.27	-0.11
75~79세	-1.91	-0.79	-2.64
80~84세	-0.61	-0.07	-1.04
85~89세	-0.70	-0.56	-0.85
90~94세	-0.22	-0.29	-0.11
95~99세	-0.05	-0.19	0.08
100세 이상	0.01	0.00	0.01

〈대 구〉

(단위 : %)

연령	전체	성별	
		남자	여자
15~19세	1.01	1.62	0.41
20~24세	4.31	5.76	2.93
25~29세	2.73	2.73	2.68
30~34세	3.34	1.67	4.84
35~39세	-3.81	-2.21	-5.26
40~44세	-0.56	-2.88	1.55
45~49세	0.81	0.12	1.44
50~54세	1.21	1.52	0.94
55~59세	-2.04	-1.81	-2.24
60~64세	-0.06	-0.48	0.34
65~69세	-3.04	-1.87	-4.09
70~74세	-1.99	-2.47	-1.54
75~79세	-0.84	-0.98	-0.69
80~84세	-0.74	-1.25	-0.25
85~89세	-0.01	0.40	-0.35
90~94세	-0.36	0.10	-0.77
95~99세	0.04	0.02	0.07
100세 이상	0.00	0.00	0.01

〈인 천〉

(단위 : %)

연령	전체	성별	
		남자	여자
15~19세	0.50	-1.10	2.10
20~24세	2.55	3.67	1.43
25~29세	3.77	4.10	3.43
30~34세	2.90	2.27	3.55
35~39세	-3.20	-2.66	-3.76
40~44세	-3.26	-2.68	-3.84
45~49세	1.38	0.61	2.16
50~54세	2.50	2.57	2.44
55~59세	-0.52	0.06	-1.11
60~64세	-2.28	-2.92	-1.64
65~69세	-1.28	-0.46	-2.09
70~74세	-0.99	-1.43	-0.56
75~79세	-1.27	-1.39	-1.14
80~84세	-0.16	-0.18	-0.13
85~89세	-0.37	-0.34	-0.40
90~94세	-0.34	-0.13	-0.56
95~99세	0.05	0.02	0.09
100세 이상	0.01	0.00	0.01

〈광 주〉

(단위 : %)

연령	전체	성별	
		남자	여자
15~19세	2.58	3.35	1.88
20~24세	5.87	9.07	3.16
25~29세	4.75	4.59	4.85
30~34세	3.34	4.57	2.31
35~39세	-0.82	-1.88	0.02
40~44세	-2.33	-5.02	-0.18
45~49세	0.14	0.76	-0.38
50~54세	2.78	0.19	4.86
55~59세	1.20	3.40	-0.60
60~64세	-2.19	0.12	-4.03
65~69세	-5.58	-8.26	-3.36
70~74세	-1.92	-2.81	-1.16
75~79세	-5.43	-4.46	-6.16
80~84세	-1.95	-2.84	-1.14
85~89세	-0.46	-0.24	-0.58
90~94세	0.25	0.11	0.39
95~99세	-0.25	-0.64	0.08
100세 이상	0.01	0.00	0.01

〈대 전〉

(단위 : %)

연령	전체	성별	
		남자	여자
15~19세	0.39	-0.74	1.76
20~24세	4.48	3.35	5.84
25~29세	3.29	1.62	5.27
30~34세	1.41	1.13	1.77
35~39세	-1.75	0.43	-4.25
40~44세	-1.64	-1.14	-2.21
45~49세	0.67	0.23	1.17
50~54세	0.69	3.04	-2.01
55~59세	0.14	0.47	-0.23
60~64세	0.09	0.64	-0.54
65~69세	-3.36	-3.51	-3.20
70~74세	-0.83	-1.65	0.04
75~79세	-2.64	-1.99	-3.45
80~84세	-1.66	-2.44	-0.86
85~89세	0.71	0.40	1.03
90~94세	-0.02	0.11	-0.20
95~99세	0.05	0.02	0.07
100세 이상	0.01	0.00	0.01

〈울 산〉

(단위 : %)

연령	전체	성별	
		남자	여자
15~19세	1.20	2.88	-0.42
20~24세	1.13	2.39	-0.19
25~29세	-0.34	0.15	-0.92
30~34세	-4.38	-6.16	-2.74
35~39세	-4.29	-4.56	-4.05
40~44세	1.35	1.04	1.68
45~49세	4.32	3.76	4.88
50~54세	2.90	3.20	2.60
55~59세	0.82	1.25	0.43
60~64세	-0.84	-2.06	0.32
65~69세	-0.04	0.03	-0.09
70~74세	-1.03	-0.97	-1.04
75~79세	-0.22	0.01	-0.37
80~84세	-0.06	-0.77	0.66
85~89세	-0.22	0.21	-0.60
90~94세	-0.32	-0.42	-0.20
95~99세	0.02	0.01	0.03
100세 이상	0.00	0.00	0.00

〈세 종〉

(단위 : %)

연령	전체	성별	
		남자	여자
15~19세	6.12	6.07	6.17
20~24세	5.93	10.60	1.59
25~29세	-0.21	-8.55	7.25
30~34세	-3.01	9.85	-14.78
35~39세	-1.99	-7.46	3.07
40~44세	2.76	-6.11	10.81
45~49세	4.70	9.77	-0.05
50~54세	-0.59	-0.64	-0.59
55~59세	-10.30	-10.90	-9.79
60~64세	-3.57	-13.23	5.34
65~69세	3.94	3.61	4.30
70~74세	-1.05	2.80	-4.48
75~79세	-5.81	2.29	-13.12
80~84세	1.87	1.32	2.48
85~89세	0.87	0.48	1.30
90~94세	0.26	0.11	0.42
95~99세	0.05	0.01	0.09
100세 이상	0.00	0.00	0.01

〈경 기〉

(단위 : %)

연령	전체	성별	
		남자	여자
15~19세	0.79	0.91	0.64
20~24세	2.33	3.46	1.35
25~29세	2.09	2.36	1.90
30~34세	1.79	1.65	1.72
35~39세	-1.54	-1.45	-2.08
40~44세	-0.69	-1.67	0.21
45~49세	-0.48	-1.55	0.63
50~54세	0.13	0.42	0.13
55~59세	-0.13	-0.17	-0.13
60~64세	-0.09	-0.50	0.21
65~69세	-1.29	-1.05	-1.33
70~74세	-0.97	-1.32	-0.55
75~79세	-1.21	-0.49	-1.97
80~84세	-0.49	-0.52	-0.46
85~89세	-0.20	-0.04	-0.25
90~94세	-0.04	-0.07	0.03
95~99세	-0.01	0.02	-0.04
100세 이상	0.01	0.00	0.01

〈강 원〉

(단위 : %)

연령	전체	성별	
		남자	여자
15~19세	1.05	1.83	0.31
20~24세	3.80	5.37	2.19
25~29세	2.52	2.49	2.35
30~34세	2.28	4.30	0.34
35~39세	-1.57	-1.35	-1.94
40~44세	0.66	0.82	0.81
45~49세	0.83	1.26	0.73
50~54세	-2.02	-4.12	0.08
55~59세	3.21	2.20	3.98
60~64세	-4.26	-2.82	-5.25
65~69세	-1.93	-2.03	-1.95
70~74세	0.52	-1.91	2.66
75~79세	-2.54	-3.34	-1.87
80~84세	-1.30	-0.81	-1.72
85~89세	-0.75	-0.92	-0.57
90~94세	-0.57	-0.88	-0.29
95~99세	0.09	0.03	0.14
100세 이상	0.01	0.00	0.01

〈충 북〉

(단위 : %)

연령	전체	성별	
		남자	여자
15~19세	3.01	2.21	3.72
20~24세	5.47	7.56	3.45
25~29세	1.50	0.36	2.51
30~34세	1.30	0.81	1.72
35~39세	-1.54	-1.88	-1.25
40~44세	1.44	2.31	0.59
45~49세	2.12	2.27	1.96
50~54세	0.96	-0.26	2.07
55~59세	0.32	0.35	0.28
60~64세	-2.17	-3.31	-1.10
65~69세	-2.37	-3.33	-1.46
70~74세	-2.73	-2.23	-3.16
75~79세	-3.19	-2.40	-3.88
80~84세	-1.66	-1.49	-1.76
85~89세	-1.31	-0.73	-1.80
90~94세	-0.97	-0.28	-1.59
95~99세	0.07	0.02	0.11
100세 이상	-0.21	0.00	-0.42

〈충 남〉

(단위 : %)

연령	전체	성별	
		남자	여자
15~19세	0.96	2.47	-0.31
20~24세	3.35	4.49	2.35
25~29세	3.41	5.52	1.56
30~34세	0.74	-0.53	1.59
35~39세	-1.48	-0.08	-2.73
40~44세	-0.87	-3.86	1.48
45~49세	2.71	2.44	2.83
50~54세	1.20	0.32	1.84
55~59세	1.39	1.40	1.37
60~64세	-0.14	-0.52	0.21
65~69세	-2.32	-0.87	-3.45
70~74세	-3.54	-4.05	-2.98
75~79세	-2.65	-4.18	-1.20
80~84세	-1.32	-1.79	-0.74
85~89세	-1.23	-0.66	-1.58
90~94세	-0.16	-0.14	-0.13
95~99세	0.10	0.03	0.16
100세 이상	-0.15	0.00	-0.28

〈전 북〉

(단위 : %)

연령	전체	성별	
		남자	여자
15~19세	0.69	-0.87	1.97
20~24세	5.01	5.32	4.64
25~29세	1.00	1.37	0.60
30~34세	0.08	0.35	-0.18
35~39세	-1.60	-1.05	-2.11
40~44세	-0.38	-0.76	-0.10
45~49세	1.67	-0.16	3.21
50~54세	3.13	2.54	3.61
55~59세	2.28	2.31	2.26
60~64세	0.09	0.92	-0.61
65~69세	-2.70	0.48	-5.40
70~74세	-3.23	-5.46	-1.24
75~79세	-2.17	-2.88	-1.46
80~84세	-3.45	-2.33	-4.30
85~89세	-0.44	0.00	-0.76
90~94세	0.10	0.20	0.04
95~99세	-0.09	0.03	-0.19
100세 이상	0.01	0.00	0.02

〈전 남〉

(단위 : %)

연령	전체	성별	
		남자	여자
15~19세	3.04	3.88	2.30
20~24세	1.42	0.16	2.45
25~29세	-0.09	1.55	-1.58
30~34세	0.31	-0.19	0.23
35~39세	-0.11	-1.51	0.57
40~44세	0.59	-0.26	0.19
45~49세	2.21	3.59	0.72
50~54세	0.96	2.10	-0.22
55~59세	0.84	-1.92	2.96
60~64세	0.64	2.77	-0.92
65~69세	-3.10	-4.51	-1.46
70~74세	-2.79	-3.81	-1.36
75~79세	-1.63	-0.21	-2.01
80~84세	-1.55	-1.69	-0.86
85~89세	-0.26	-0.29	0.05
90~94세	-0.32	0.28	-0.63
95~99세	-0.25	0.04	-0.44
100세 이상	0.02	0.00	0.03

〈경 북〉

(단위 : %)

연령	전체	성별	
		남자	여자
15~19세	0.19	-1.65	1.86
20~24세	3.23	3.49	2.93
25~29세	0.17	0.71	-0.41
30~34세	0.44	1.55	-0.63
35~39세	-1.22	-2.13	-0.41
40~44세	-1.15	-0.63	-1.65
45~49세	1.30	1.43	1.15
50~54세	1.51	1.99	1.05
55~59세	0.25	-1.23	1.61
60~64세	0.56	1.14	0.03
65~69세	-1.73	-0.93	-2.46
70~74세	-1.54	-2.64	-0.47
75~79세	-2.09	-0.58	-3.42
80~84세	-0.44	-1.07	0.22
85~89세	0.17	0.26	0.14
90~94세	0.37	0.22	0.53
95~99세	-0.03	0.04	-0.08
100세 이상	0.01	0.00	0.02

〈경 남〉

(단위 : %)

연령	전체	성별	
		남자	여자
15~19세	1.92	0.64	3.14
20~24세	2.34	3.90	0.77
25~29세	1.12	1.62	0.59
30~34세	-0.20	0.25	-0.67
35~39세	-2.64	-2.74	-2.57
40~44세	-0.49	-0.10	-0.88
45~49세	1.23	-0.41	2.82
50~54세	1.35	1.39	1.32
55~59세	2.27	2.07	2.48
60~64세	-1.89	-0.70	-3.04
65~69세	-2.62	-2.75	-2.47
70~74세	-0.69	-1.20	-0.17
75~79세	-1.49	-1.47	-1.46
80~84세	-0.02	-0.20	0.21
85~89세	-0.04	0.14	-0.19
90~94세	-0.21	-0.43	0.02
95~99세	0.06	0.02	0.10
100세 이상	0.01	0.00	0.01

〈제 주〉

(단위 : %)

연령	전체	성별	
		남자	여자
15~19세	-2.58	-3.39	-1.84
20~24세	7.20	8.71	5.76
25~29세	4.20	4.61	3.79
30~34세	3.43	6.17	0.85
35~39세	1.79	0.07	3.39
40~44세	3.63	3.72	3.51
45~49세	0.84	0.10	1.50
50~54세	3.65	2.92	4.30
55~59세	-6.04	-0.19	-11.51
60~64세	-2.62	-6.52	1.01
65~69세	0.74	3.06	-1.41
70~74세	-4.21	-6.69	-1.88
75~79세	-4.17	-5.97	-2.44
80~84세	-2.64	-1.96	-3.23
85~89세	-3.97	-4.81	-3.15
90~94세	0.57	0.14	1.00
95~99세	0.14	0.03	0.26
100세 이상	0.03	0.00	0.06

〈부표 2〉 지역별 가구 유형에 대한 추계자료와 패널자료의 비율 차이

(단위 : %)

	가구원수					
	1인	2인	3인	4인	5인	6인 이상
전 국	3.99	2.15	1.61	-6.96	-0.89	0.10
서 울	3.68	0.86	0.03	-5.12	-0.03	0.59
부 산	0.82	3.76	1.55	-6.16	-0.10	0.12
대 구	2.09	3.87	2.76	-6.32	-2.88	0.49
인 천	4.34	0.65	3.69	-7.28	-1.09	-0.31
광 주	-0.59	-1.05	2.88	-0.30	-0.59	-0.35
대 전	5.52	1.51	-1.47	-6.58	-0.33	1.36
울 산	3.05	9.96	1.01	-13.91	-0.15	0.03
세 종	2.04	2.76	0.87	-13.59	6.06	1.88
경 기	5.38	3.07	1.56	-9.04	-0.83	-0.15
강 원	8.69	2.85	-0.54	-6.88	-2.88	-1.23
충 북	4.40	2.22	1.87	-9.71	0.34	0.88
충 남	2.09	3.94	2.82	-3.38	-4.74	-0.72
전 북	3.24	1.72	1.72	-6.57	0.70	-0.82
전 남	5.82	-1.97	1.95	-6.87	0.07	1.00
경 북	4.55	2.57	0.94	-6.47	-1.28	-0.30
경 남	4.90	2.66	3.72	-9.72	-1.72	0.16
제 주	3.63	-6.79	3.40	1.58	-2.47	0.64

제 3 장
KILPS 보육관련 설문조사의 문제와 데이터 품질 제고 방안

제1절 문제 제기

노동패널조사에는 가구 내 자녀에 대한 교육・보육서비스 이용 실태에 관한 문항이 다수 포함되어 있다. 가구 내 고등학생 이하 자녀에 대한 교육・보육서비스 이용에 관한 설문 문항은 2차년도와 6차년도를 제외하고는 매해 조사되는 항목이지만, 조사 문항과 설문의 구조는 지속적으로 변화 발전해 왔다. 보육관련 설문 문항은 5차년도 이전까지 주로 서비스의 종류를 수정 보완하는 형태로 이루어졌으며, 6차년도 조사부터 15차년도 조사까지는 설문에 큰 변화를 주지 않고 진행되었다. 다만, 12차년도(2009년)에 노동패널조사 표본 추가가 있었기 때문에, 영유아 가구의 수 또한 12차년도(2009년) 조사부터 표본의 수가 크게 증가하는 특징이 있다. 이후 16차년도(2013년) 조사부터 응답자 편의를 고려하여 설문 문항은 영유아기와 학령기(초등학교~고등학교)[3] 자녀로 분리하여 질문하였다.

그런데, 이와 같이 설문 구조와 문항을 수정하는 과정에서 보육연령에 대한 CAPI 로직상의 오류가 발생하여, 16차년도(2013년) 조사 이후 보육연령 5세(취학 전 1년)의 어린이집 및 유치원 이용 실태 및 이용비용 문항에 대한 조사가 누락되는 현상이 발생하였다. 이러한 조사상 오류가 가

3) 대학생 이상 자녀의 교육비에 대한 설문은 16차년도 이전에도 별도로 존재한다.

구 내 일부 연령의 자녀에 국한하여 발생함에 따라 오류의 발견이 늦어지게 되어 19차(2016년) 조사 초반에서야 오류 수정의 기회를 갖게 되었다. 즉 결론적으로 16차년도(2013년조사)~18차년도(2015년조사) 자료는 보육연령 5세(취학 전 1년)의 어린이집 및 유치원 이용 실태 및 이용비용이 전체적으로 누락되었으며, 19차년도 조사 초기 응답 가구 중 일부에서 응답 값이 없는 상황이 발생하게 되었다.

노동패널조사팀은 오류를 발견한 즉시 조치를 취하여, 19차(2016년) 조사에서 설문 로직을 수정하여 조사를 실시하였으며, 16차년도(2013년조사)~18차년도(2015년조사) 미응답자에 대해 보육연령 5세 당시 상황에 대한 회고조사를 실시했다. 참고로, 보육연령 5세에 대한 조사가 누락되기 시작한 2013년도는 영유아 가구에 대한 무상 유아교육·보육 정책이 소득에 무관하게 전 가구 전 연령의 영유아에게로 확장된 시기로, 이 시기를 전후한 영유아 가구의 교육·보육서비스 이용 행태 및 이용비용의 변화는 무상 유아교육·보육 정책의 효과 분석 등의 연구에 반드시 필요한 자료이다. 또한, 무엇보다 노동패널조사 자료의 영유아 보육·교육 관련 설문 문항은 가구 조사 설문에서 아동 단위 응답 값을 취합하여 생성되는 형태이기 때문에, 특정 연도에 특정 연령의 자녀에 한해서 비용이 누락되는 현상은 전반적인 영유아 가구의 교육·보육서비스 이용 및 이용비용에 대한 왜곡을 초래하게 된다. 이에 노동패널팀은 19차(2016년) 조사에서 16차(2013년)~18차(2015년) 조사 당시 보육연령 5세 자녀가 있는 가구를 대상으로 보육·교육서비스 이용에 관한 회고조사를 별도로 실시하여 자료를 보완한 것이다. 다만, 해당 연령의 자녀에 대한 정보를 기존 자료와 함께 취합하여 제공하기 위해서는 회고조사 자료의 응답이 어느 정도 신뢰할 수 있는지에 대한 검토가 필요하다.

따라서 본 연구는 노동패널 16차~19차 자료의 영유아 보육·교육서비스 회고조사 자료의 특징과 이를 원자료에 포함시켜 분석할 경우 발생할 수 있는 오류를 사전에 진단하고, 응답 값의 신뢰도를 검토하는 것을 주요 목적으로 한다. 다만, 해당 오류가 취학 1년 전 어린이집 및 유치원 이용비용이 전반적으로 누락됨에 따라 발생한 것이기 때문에, 동일 연령의 아이 중에서 회고조사와 아닌 경우를 비교하기 어렵고, 자녀 연령 증가에

따라 서비스 이용의 변화가 급격하기 때문에 전년도 자료와의 비교도 쉽지 않은 상황이다. 즉 보육·교육서비스의 이용은 아동의 연령 증가 및 정책 변화 등에 따라 변화하는 것이 당연한 것이기 때문에, 다른 연령 및 다른 차수와의 비교, 종단자료로서 코호트 간의 비교 등을 통해 전반적인 추이에 이상한 변화가 섞여 있지는 않은지를 검증하는 수준에서 자료의 특징을 살펴보고 있음을 주지해 주기 바란다.

제2절 설문의 구조 변화와 누락 사유

먼저, 노동패널조사의 영유아 보육·교육서비스 이용에 관한 설문 구조의 변화와 문항에 대해 살펴본다. 16차년도 이후 설문에 있어서 가장 큰 변화는, 고등학생 이하 자녀에 대해서 묻는 것을 영유아(취학 전) 자녀와 학령기(초등학교~고등학교) 자녀로 구분하여 질문함으로써, 해당 연령에 해당하는 서비스 항목만 구분하여 질문하였다는 차이가 있다. 또한, 이 과정에서 영유아 자녀의 경우에는 어린이집 혹은 유치원과 같은 기관 이용비용과 아이돌봄서비스 이용비용만 조사되었고, 사교육 관련 문항을 질문하지 않은 차이가 있다. 이때, 보육연령 5세(취학 1년 전) 자녀의 경우 CAPI 설계상에서의 로직 오류로 0세~미취학 자녀 문항에서 응답되지 않고 일부 응답자가 사교육 서비스의 일부에 응답되는 경우가 발생하였다. 하지만, 16차년도 이후 영유아 자녀의 경우 전체적으로 사교육 서비스 이용에 관한 문항을 묻지 않았기 때문에 일관성을 확보하기 위해 애초 조사 당시 응답한 보육연령 5세의 사교육 문항은 자료에서 삭제하였다. 한편 학령기 자녀의 경우에는 어린이집이나 유치원 이용은 하지 않으므로 상관없지만, 초등 저학년의 이용 빈도가 높은 개별 돌봄서비스 등을 질문하지 않은 차이가 있다. 각 항목별로 설문의 변화를 살펴보면, <표 3-1>과 같다.

한편, 이처럼 변경된 노동패널조사 자료의 보육·교육서비스 종류에 관한 설문 문항의 개선점을 간략히 정리하면, <표 3-2>와 같다. 이와 같

〈표 3-1〉 미취학아동의 교육·보육 서비스 이용 설문 구조의 변화

<table>
<tr><th colspan="2" rowspan="3">응답차수</th><th colspan="6">설문 문항 번호(가구용 설문)</th></tr>
<tr><th>15차 이전</th><th colspan="3">16차~18차</th><th colspan="2">19차~20차</th></tr>
<tr><th colspan="2" rowspan="2">응답대상</th><th rowspan="2">0세~고등학생 이하 자녀</th><th colspan="2">0세~미취학</th><th rowspan="2">취학~고등학생 이하 자녀</th><th rowspan="2">0세~미취학 자녀</th><th rowspan="2">취학~고등학생 이하 자녀</th></tr>
<tr><th>기관 서비스 이용자</th><th>기관 서비스 미이용자</th></tr>
<tr><td rowspan="15">교육·보육서비스 종류</td><td>자녀의 부모</td><td>항목없음</td><td rowspan="4">라 7-6</td><td rowspan="4">라 7-7</td><td rowspan="4">조사안함</td><td rowspan="4">라 7-6 혹은 라 7-7</td><td rowspan="4">조사안함</td></tr>
<tr><td>동거하고 있는 가족과 친지</td><td rowspan="12">라 3-1</td></tr>
<tr><td>비동거하고 있는 가족과 친지</td></tr>
<tr><td>동거하고 있는 타인(아이 보는 아줌마, 파출부 등)</td></tr>
<tr><td>비동거하고 있는 타인(아이 보는 아줌마, 파출부 등)</td><td colspan="2" rowspan="3">로직오류로 일부 조사되었으나, 자료 삭제</td><td rowspan="3">라 8-3</td><td rowspan="3">조사안함</td><td rowspan="3">라 8-3</td></tr>
<tr><td>학원</td></tr>
<tr><td>개인·그룹과외</td></tr>
<tr><td>학습지</td><td colspan="2" rowspan="2">조사 불필요</td><td rowspan="2">라 8-3 (설문 문구 일부 수정)</td><td rowspan="2">조사 불필요</td><td rowspan="2">라 8-3</td></tr>
<tr><td>방과후 교내 보충학습(만 6세 이상으로 학교 내에서 자발적으로 별도의 비용을 지불하고 보충학습을 하는 경우)</td></tr>
<tr><td>방과후 교내 교육 및 보육(만 6세 이상 12세 이하 초등학생 자녀 : 방과후교실, 방과후 돌봄 교실 등 학교 내에서 서비스 이용)</td><td colspan="2" rowspan="3">조사 불필요</td><td rowspan="3">라 8-3 (설문 문구 일부 수정)</td><td rowspan="3">조사 불필요</td><td rowspan="3">라 8-3</td></tr>
<tr><td>방과후 교실(만 6세 이상으로 학교가 아닌 사회복지관, 유치원 등에서 보육이 아닌 특기지도, 보충학습 등을 하는 것)</td></tr>
<tr><td>방과후 학교 이외 교육 및 보육(만 6세 이상 12세 이하 초등학생 자녀 : 학교가 아닌 시설에서 서비스 이용 - 사회복지관, 지역아동센터, 청소년 수련관 및 문화의 집 등)</td></tr>
</table>

〈표 3-1〉의 계속

<table>
<tr><th colspan="2" rowspan="3">응답차수</th><th colspan="6">설문 문항 번호(가구용 설문)</th></tr>
<tr><th>15차 이전</th><th colspan="3">16차~18차</th><th colspan="2">19차~20차</th></tr>
<tr><th colspan="2" rowspan="2">응답대상</th><th rowspan="2">0세~고등학생 이하 자녀</th><th colspan="2">0세~미취학</th><th rowspan="2">취학~고등학생 이하 자녀</th><th rowspan="2">0세~미취학 자녀</th><th rowspan="2">취학~고등학생 이하 자녀</th></tr>
<tr><th>기관 서비스 이용자</th><th>기관 서비스 미이용자</th></tr>
<tr><td rowspan="10">교육·보육서비스 종류</td><td>국공립(국립·사립·구립 등) 어린이집</td><td rowspan="10"></td><td colspan="2" rowspan="7">라 7-3</td><td rowspan="7">조사 불필요</td><td rowspan="7">라 7-3</td><td rowspan="7">조사 불필요</td></tr>
<tr><td>민간 어린이집</td></tr>
<tr><td>사설놀이방(개인 가정에서 운영하는 놀이방)</td></tr>
<tr><td>직장보육시설</td></tr>
<tr><td>정규유치원(4시간)</td></tr>
<tr><td>시간연장제 유치원(6시간)</td></tr>
<tr><td>종일제 유치원(8시간)</td></tr>
<tr><td>어학연수(영어캠프 포함)</td><td colspan="2" rowspan="3">로직오류로 일부 조사되었으나, 자료 삭제</td><td rowspan="3">라 8-3</td><td rowspan="3">조사안함</td><td rowspan="3">라 8-3</td></tr>
<tr><td>인터넷(온라인 교육) 유료 강좌(EBS TV 등의 강좌 포함)</td></tr>
<tr><td>문화센터</td></tr>
<tr><td>기타</td><td></td><td colspan="2"></td><td></td><td></td><td></td></tr>
<tr><td>종일제 아이돌보미서비스</td><td>기타에 포함</td><td colspan="2">라 7-3</td><td>조사안함</td><td>라 7-3</td><td>조사안함</td></tr>
</table>

〈표 3-2〉 교육·보육 서비스 종류 문항의 용어 및 설문 수정 고려 사항(19차 설문 기준)

<table>
<tr><th colspan="2">19차 설문 기준</th><th>0세~미취학 자녀</th><th>취학~고등학생 이하 자녀</th></tr>
<tr><td rowspan="4">교육·보육서비스 종류</td><td>자녀의 부모</td><td rowspan="4">개별돌봄서비스 활용을 물어보는데, 여성가족부의 종일제 아이돌봄서비스는 라7-3에서 기관서비스와 함께 묻고 있음.</td><td rowspan="4">- 학령기 자녀의 경우(특히 초등 저학년 자녀)에도 돌봄서비스를 이용하는 경우가 다수 있음.
- 학령기 자녀의 돌봄서비스 이용을 묻지 않는 문제가 있음.</td></tr>
<tr><td>동거하고 있는 가족과 친지</td></tr>
<tr><td>비동거하고 있는 가족과 친지</td></tr>
<tr><td>동거하고 있는 타인(아이 보는 아줌마, 파출부 등)
비동거하고 있는 타인(아이 보는 아줌마, 파출부 등)</td></tr>
</table>

〈표 3-2〉의 계속

19차 설문 기준		0세~미취학 자녀	취학~고등학생 이하 자녀
교육·보육서비스 종류	학원	사교육 문항은 16차부터 조사하지 않음. 향후에도 영유아 사교육은 조사하지 않을 것인지 결정 필요	현행 유지
	개인·그룹과외		
	학습지		
	방과후 교내 교육 및 보육(만 6세 이상 12세 이하 초등학생자녀 : 방과후교실, 방과후 돌봄 교실 등 학교 내에서 서비스 이용)	"방과후 학교"는 초등 이상 자녀만 조사하는 것이 맞음. 유치원의 '방과후 학교'와 특성화프로그램, 어린이집 특별활동 비용이 조사되어야 함. 아래 유치원, 어린이집, 이용비용에 특성화프로그램 및 특별활동 비용 등이 포함되는지 명확한 가이드 필요	"방과후 학교"가 정확한 용어임. 초등돌봄을 포함한다는 것인지 모호함.
	방과후 학교 이외 교육 및 보육(만 6세 이상 12세 이하 초등학생 자녀: 학교가 아닌 시설에서 서비스 이용 - 사회복지관, 지역아동센터, 청소년 수련관 및 문화의 집 등)		용어정의가 모호함. 서비스 제공기관이 다름을 강조하여 수정할 필요가 있음
	국공립(국립·사립·구립 등) 어린이집	어린이집은 설립유형으로 구분되어 있는데, 어린이집 설립유형별 구분은 총 7개임. 아래 직장어린이집까지 3개 설립유형이 제시되어 있고, 4개(법인/법인·단체/부모협동/가정어린이)는 응답할 곳이 없음 특히 개소수가 가장 많은 가정어린이집의 구분이 없는 것은 큰 문제임.	현행 유지 (조사필요없음)
	민간 어린이집		
	사설놀이방(개인 가정에서 운영하는 놀이방)	이 문항이 "가정어린이집"? 사설놀이방은 기본적으로 어린이집이 아닌 반일제 학원 등임. 용어가 명확하게 구분되어야 함.	
	직장보육시설	"직장 어린이집"이 정확한 명칭이며, 설립유형 중 하나임.	
	정규유치원(4시간)	유치원의 경우에는 이용시간을 기준으로 구분하고 있음. 설립유형 기준 국공립/사립 유치원으로 구분됨.	
	시간연장제 유치원(6시간)		
	종일제 유치원(8시간)		
	어학연수(영어캠프 포함)	사교육 문항은 16차부터 조사하지 않음. 향후에도 영유아 사교육은 조사하지 않을 것인지 결정 필요	현행 유지
	인터넷(온라인 교육) 유료 강좌(EBS TV 등의 강좌 포함)		
	문화센터		
	기타	현행 유지	
	종일제 아이돌보미서비스	(라-6과 라-7로 미루어보아) 설문구조상 여성가족부의 아이돌보미서비스를 의미하는 것으로 보이며, 명확하게 표기	여성가족부 돌봄서비스도 초등학령기까지를 지원대상으로 하고 있음.

은 설문 개선에 대한 의견은 노동패널조사 자료가 오랜 기간에 걸쳐 조사되는 과정에 종단면 연속성 확보를 위해 설문 항목(서비스 종류)을 크게 고치지 않았기 때문에 발생한 문제로, 부정확한 분류 기준 혹은 정책 용어의 사용으로 인해 조사 과정에서의 혼선과 응답 값에 대한 오해석 위험이 높아졌기 때문이다.

16차년도(2013년) 조사부터는 미취학 자녀의 사교육 서비스 이용 문항이 삭제된 것에 반해, 2013년부터 시작된 무상 보육・유아교육에 대한 설문 문항(아이사랑 카드, 가족양육수당) 등 정책 수혜 여부와 사교육비 부담 등에 관한 설문 문항이 추가되었다. 추가된 설문 문항과 2017년 기준으로 살펴보았을 때 이들 설문의 개선점을 정리하면 <표 3-3>과 같다.

한편, 회고조사 자료의 값을 검증하기 위해서 자료를 몇 가지 카테고리로 분류하여 사용하였다. 현재 교육・보육서비스 종류가 너무 많아 연령

〈표 3-3〉 교육・보육비용 지원 정책 관련 설문 문항 및 개선점(16차 이후 추가 문항)

16차년도 이후 추가 설문 문항		설문문항 번호 (가구용 설문)	설문 개선 필요사항
0세~미취학 자녀 대상 설문 문항	아이사랑카드, 아이즐거운카드 사용여부	라 7-9	- 2015년 1월부터 "아이행복카드"로 통합하여 운영되고 있음. - 바우처 지원은 카드를 통해서만 시행되므로 기관 이용 여부와 사용 여부는 중복문항임(해당 카드는 신용카드이므로, 다른 용도로 사용하느냐를 묻는 것으로 오인될 소지가 있음).
	자녀 양육수당 수혜 여부	라 7-9	- "가정양육수당"이 정확한 정책 명칭임. - 2018년 7월부터 "아동수당" 지원이 신설되므로, 오인의 소지가 높음.
	지급받는 양육수당 액수	라 7-10	"가정양육수당" 액수는 월령에 따라 정해진 액수가 있으며, 이를 응답 가구가 월평균 하여 응답하려면 기준 시점이 주어져야 함.
	사교육비로 인한 부담	라 7-11	- 사교육비 문항자체가 없으며, 정의가 모호 - 어린이집, 유치원 이용비용을 사교육비라고 볼 수는 없으므로, 사교육에 대한 범주를 예시로라도 제시해야 함.

〈표 3-3〉의 계속

16차년도 이후 추가 설문 문항		설문문항 번호 (가구용 설문)	설문 개선 필요사항
0세~미취학 자녀 대상 설문 문항	보육 및 자녀 양육비 지원이 사교육 부담 경감에 주는 도움 정도	라 7-12	- "보육 및 자녀 양육비 지원"의 범주도 모호함. - 어린이집 혹은 유치원 바우처는 "유아학비·보육료 지원" 혹은 "누리과정지원금 및 보육료 지원금"이 정확한 표현이며, 가정에 주는 현금수당은 "가정양육수당" 지원이 정책 용어임. 세 가지 지원 정책을 모두 포괄하여 표현하고자 한다면, "무상 유아교육·보육" 등으로 표현하거나, "유아학비·보육료 및 가정양육수당"으로 표기하는 것이 바람직함. - 2018년 하반기부터 "아동수당"지원이 추가되기 때문에, 향후에는 정확하게 어떤 정책에 대한 선호인지를 구분하는 것이 더 중요한 사항임.
취학~고등학생 이하 자녀(학령기 자녀) 대상 설문 문항	방과후 교내 교육 및 보육서비스 비용 지원 여부	라 8-5	- "방과후 학교", "초등돌봄교실" 등이 정책 명칭으로, 2개의 서비스는 다른 서비스임(현재 '방과후 학교 연계형 돌봄 교실' 연구가 교육개발원에서 수행 중이기는 함). - 서비스 지원 기관이 학교인가 아닌가를 기준으로 한 구분으로 보이므로, "학교에서 실시하는 방과후 학교 및 초등돌봄교실", "지역사회(지역아동센터, 사회복지관, 청소년수련관 등) 연계형 방과후 학교 및 초등돌봄교실"로 명확하게 표기하는 것이 응답하기에 용이함.
	방과후 학교 이외 교육 및 보육		
	사교육비로 인한 부담	라 8-6	- 사교육의 범주를 명확히 해야 함. - '방과후 학교' 등을 사교육에 포함할 것인지 말 것인지는 연구자에 따라 상이함. 참고로 교육부는 방과후 학교 등을 사교육비가 아닌 사부담 공교육비로 분류함.
	보육 및 자녀 양육비 지원이 사교육 부담 경감에 주는 도움 정도	라 8-7	- 어떠한 정책에 대한 도움 정도를 묻는 것인지 모호함. - 학령기 자녀의 방과후 학교 지원은 현재로서는 저소득층 등 취약계층에 국한한 지원 사업임. - '초등돌봄교실' 혹은 '방과후 학교'가 이에 해당하는지 여부도 모호함.

별, 각 서비스별 이용자수가 너무 적고, 서비스 종류의 구분이 크게 의미가 없는 경우를 기준으로 하여 정리한 것이다. 본 연구에서 사용한 분류기준과 해당되는 항목은 <표 3-4>와 같다. 사선으로 표시된 서비스와 이용비용이 15차년도 이전과 16차년도 이후가 연속성을 갖는 변수들이다.

〈표 3-4〉 코드값 변환

본 연구 분류 기준	노동패널 16차 이후 설문 항목 기준	특기사항
어린이집	국공립(국립·사립·구립 등) 어린이집 민간 어린이집 직장보육시설	
2. 유치원	정규유치원(4시간) 시간연장제 유치원(6시간) 종일제 유치원(8시간)	
3. 개별돌봄 서비스	종일제 아이돌보미서비스(여성가족부 사업)	16차 이후 추가된 항목으로 어린이집, 유치원 이용과 함께 질문하고 있음
	동거하고 있는 가족과 친지 비동거하고 있는 가족과 친지 동거하고 있는 타인(아이보는 아줌마, 파출부 등) 비동거하고 있는 타인(아이보는 아줌마, 파출부 등)	
4. 학원	학원 사설놀이방(개인가정에서 운영하는 놀이방) 문화센터	해당 문항은 사교육비로, 16차 조사 이후 미취학 아동에 대해서는 질문하지 않음.
5. 학습지	학습지	
6. 개인교습	개인·그룹과외	
7. 인터넷	인터넷(온라인 교육) 유료 강좌(EBS TV 등의 강좌 포함)	
8. 부모 직접 돌봄(가정 내 양육)	자녀의 부모	- 16차 이후 추가된 문항으로, 15차년도와 분류방식이 다소 차이가 있음. - 15차 이전 자료는 어떠한 서비스도 이용하지 않는 경우에 한해서, 이 경우로 분류가 가능함.

한편, 분류된 서비스 중에서도 4~7번까지는 사교육비에 해당하는 서비스로, 16차년도 조사 이후로는 사교육을 질문하지 않았으므로 추이 변화 등을 비교하지 않는다. 결론적으로 본 연구가 살펴본 서비스는 어린이집 이용 행태, 유치원 이용행태, 개별돌봄서비스 이용 행태이다.

제3절 회고조사 결과

16차년도 이후 보육연령 5세의 보육·교육서비스 설문이 누락된 자료의 수는 16차 215가구, 17차 211가구, 18차 203가구로 총 629가구이며, 19차년도 조사 초반 조사되지 않아 자료 보완이 필요한 가구가 52가구였다. 19차년도 조사는 조사 중반 로직이 수정되어, 로직 수정 전에 응답한 가구에 대한 재조사를 실시하였으므로 엄밀한 의미에서 회고조사 자료라 보기는 어렵다. 가구별로 2명 이상의 자녀에 대한 누락(보육연령 5세만 누락되었으므로 대체로 쌍둥이 가구)이 있는 경우가 있어서, 조사 완료 아동수 기준으로는 조사완료 가구수에 비해 약간 높다. 아동을 기준으로 할 경우 16차년도 212명, 17차년도 205명, 18차년도 198명이 조사되었다.

반면, 보육·교육설문에 대한 재조사는 이루어졌으나 원조사 당시 다른 변수들에 미싱이 존재하여 본 연구의 분석에 활용된 자료는 16차년도 192명, 17차년도 192명, 18차년도 185명이다. 재조사가 됐음에도 분석 자료에서 빠진 이유는 주로 자녀의 어머니가 개인용 설문에 응답하지 않았거나 모 정보가 누락된 경우였으며, 가구소득 등이 미싱이 있는 경우도 일부 존재한다. 즉 본 연구의 분석 자료는 13차년도(2010년)~19차년도(2016년)까지 취학 전 아동(보육연령 5세 이하 전부)에 대한 분석으로, 가구조사로 이뤄진 자료를 아동 단위로 변환한 것으로, 가구 정보 및 모 정보 등을 응답하지 않은 경우에는 다른 차수 자료와의 일관성 확보를 위해 분석에서 제외한다. 이때, 회고조사 자료는 16차(2013년)~18차(2015년)의 보육연령 5세에만 해당한다.

〈표 3-5〉 회고조사 자료

(보육연령 5세)		16차년도 (2013년)	17차년도 (2014년)	18차년도 (2015년)	16~18차 합계	19차년도 (2016년)
누락된 조사 자료		215	211	203	629	52
회고조사	조사 완료 가구수	209	201	195	605	50
	조사 완료 아동수	212	205	198	615	50
	본 연구의 분석 자료 (아동수)	192	192	185	569	50
재조사완료율(유효한 값 기준)		89.3	91.0	91.1	90.5	96.2

분석에 사용된 자료의 대상자수는 <표 3-6>과 같으며, 표에서 음영에 굵은 글자로 표시된 경우가 회고조사로 응답된 분석대상자의 수이다. 참고로, 보육 연령 5세의 분석대상자 수를 살펴보면, 2010년 181명, 2011년 162명, 2012년 226명, 2013과 2014년 192명, 2015년 185명, 2016년 208명이다. 2011년에는 5세 아동이 적은 편이고, 2012년에는 반대로 5세 아동수가 많은 특징이 있다. 한편, 동일 차수 내에서 연령별 아동의 수를 살펴보면, 회고조사 자료(보육연령 5세)의 비중이 16차 15.1%, 17차 16.4%, 18차 15.6%로 다른 차수들과 큰 차이를 보이지는 않았다. 표에서 해당 회고자료의 좌측 대각선 윗방향의 대상자수는 동일 코호트의 아이들 수를 의미한다. 즉 16차년도 5세 아이들은 15차년도에는 4세였을 것이므로 동일 코호트(16차 기준 2007년생)의 15차년도 분석 대상은 206명인 데 반해 16차년도 회고조사 자료에서는 192명이 응답했으므로, 14명이 줄어들었음을 알 수 있다. 16차년도 4세(203명)와 17차년도 5세(192명)는 11명이 감소했으나, 17차년도 4세(187명)와 18차년도 5세(185명) 사이에는 단 2명이 차이가 났다. 둘 다 회고조사 자료가 아닌 18차년도 4세(210명)와 19차년도 5세(208명) 사이에도 2명의 차이가 있다. 이를 통해 회고조사가 이뤄진 19차(2016년)에서 비교적 회고 시점이 먼 16차년도(2013년) 조사의 경우에 회고조사 응답률이 다소 낮음을 알 수 있으며, 조사시점부터 회고시점이 멀지 않은 경우에는 응답률이 다소 높음을 엿볼 수 있다. 이를 통해 조사상 오류가 발견되어 재조사 및 회고조사 등을 실시해야 하는 경우에는 가능한 한 빠른 시일 내에 조사 자료의 보완을 실시하는 것이 응답

률 및 자료 신빙성을 높일 수 있는 길임을 느낄 수 있었다.

회고조사가 이뤄진 연령인 보육연령 5세(취학 1년 전) 아동의 인적 특성을 성별과 출생순위별로 살펴보면, 남아가 여아에 비해 약간 많은 편이지만, 19차(2016년)에만 남아가 여아에 비해 4명 적었다. 19차년도 조사자료는 회고조사 자료가 아니므로 논외로 하고, 16차~18차 조사 자료에서는 17차(2014년) 조사 자료가 남아가 60.4%, 여아가 39.6%로 성별 비중이

〈표 3-6〉 연도(조사차수)별, 연령별 분석 대상자수

(단위 : 명, %)

		보육연령별 응답자 수						
		0세	1세	2세	3세	4세	5세	전체
기준연도	2010년(13차)	262(21.3)	191(15.5)	202(16.4)	223(18.1)	170(13.8)	181(14.7)	1,229(100.0)
	2011년(14차)	307(24.0)	183(14.3)	196(15.3)	205(16.0)	227(17.7)	162(12.7)	1,280(100.0)
	2012년(15차)	298(22.8)	192(14.7)	188(14.4)	200(15.3)	206(15.7)	226(17.3)	1,310(100.0)
	2013년(16차)	295(23.2)	194(15.3)	196(15.4)	192(15.1)	203(16.0)	192(15.1)	1,272(100.0)
	2014년(17차)	203(17.3)	197(16.8)	195(16.6)	198(16.9)	187(16.0)	192(16.4)	1,172(100.0)
	2015년(18차)	229(19.3)	163(13.8)	203(17.1)	195(16.5)	210(17.7)	185(15.6)	1,185(100.0)
	2016년(19차)	190(16.6)	175(15.3)	167(14.6)	205(18.0)	197(17.3)	208(18.2)	1,142(100.0)

주 : 음영 표시된 부분이 회고조사 자료를 의미함.

〈표 3-7〉 분석 대상자 특성(보육연령 5세만)

(단위 : 명, %)

		조사연도(차수)						
		2010년 (13차)	2011년 (14차)	2012년 (15차)	2013년 (16차)	2014년 (17차)	2015년 (18차)	2016년 (19차)
전체		181(100.0)	162(100.0)	226(100.0)	192(100.0)	192(100.0)	185(100.0)	208(100.0)
성별	남아	105(58.0)	84(51.9)	125(55.3)	98(51.0)	116(60.4)	97(52.4)	102(49.0)
	여아	76(42.0)	78(48.2)	101(44.7)	94(49.0)	76(39.6)	88(47.6)	106(50.8)
출생순위	첫째	101(55.8)	85(52.5)	107(47.4)	93(48.4)	93(48.1)	89(48.1)	106(51.0)
	둘째	72(39.8)	66(40.7)	102(45.1)	80(41.7)	78(40.6)	78(42.2)	80(38.5)
	셋째	7(3.9)	11(6.8)	16(7.1)	16(8.3)	19(9.9)	17(9.2)	21(10.1)
	넷째 이상	1(0.6)	0(0.0)	1(0.4)	3(1.6)	2(1.0)	1(0.5)	1(0.5)

주 : 음영 표시된 부분이 회고조사 자료임.

〈표 3-8〉 분석대상자 가구 특성(보육연령 5세만)

(단위 : 만 원, 명, %)

		조사연도 (차수)						
		2010년 (13차)	2011년 (14차)	2012년 (15차)	2013년 (16차)	2014년 (17차)	2015년 (18차)	2016년 (19차)
N		181(100.0)	162(100.0)	226(100.0)	192(100.0)	192(100.0)	185(100.0)	208(100.0)
가구소득 평균		350.0	382.7	378.5	414.7	408.9	446.0	470.5
가구소득	200만 원 미만	20(11.1)	19(11.7)	17(7.5)	10(5.2)	13(6.8)	13(7.0)	10(4.8)
	200~300만 원 미만	59(32.6)	35(21.6)	58(25.7)	32(16.7)	38(19.8)	28(15.1)	26(12.5)
	300~400만 원 미만	43(23.8)	39(24.1)	60(26.6)	59(30.7)	54(28.1)	47(25.4)	53(25.5)
	400~500만 원 미만	27(14.9)	28(17.7)	40(17.7)	35(18.2)	35(18.2)	37(20.0)	30(14.4)
	500~600만 원 미만	13(7.2)	18(10.2)	23(10.2	29(15.1)	22(11.5)	20(10.8)	38(18.3)
	600만 원 이상	19(10.5)	23(14.2)	28(12.4)	27(14.1)	30(15.6)	40(21.6)	51(24.5)

주 : 음영 표시된 부분이 회고조사 자료임.

다른 차수와 다소 차이를 보였다. 표에는 제시하지 않았으나, 회고조사 자료가 아닌 16차 조사 당시 4세 아동들의 성별 비중을 산출해 본 결과, 남아가 123명(60.6%), 여아가 80명(39.4%)으로 애초에 노동패널조사 자료에서 2008년생 코호트(16차 조사 당시 4세, 17차 조사 당시 5세아)가 남아의 비중이 여아의 비중에 비해 높은 경우로, 회고조사 응답자가 특정 성별에 쏠려서 더 조사됐다거나 하지는 않았음을 알 수 있었다. 출생 순위별로는 회고조사 이전과 이후 조사에서 큰 차이를 보이지 않았다.

가구소득 수준은 전반적으로 증가하는 추세를 보이는데 15차(2012년)와 17차(2014년)에는 전년대비 평균 가구소득이 약간 낮아지기는 하지만, 전반적인 가구소득의 분포를 살펴보면 크게 차이를 보이는 구간이 없는 것으로 나타났다. 즉 가구소득이 연도별로 약간의 변동은 있으나, 특정 소득 구간의 가구가 체계적으로 회고조사에서 누락되거나 이탈했다고 보기는 어려워 보인다.

제4절 보육 · 교육서비스 이용 행태 변화

1. 보육 · 교육서비스 이용 비중 변화

회고조사 자료의 보육 · 교육서비스 이용 행태를 살펴보기에 앞서 결과표 해석의 이해를 돕기 위해 아래의 <표 3-9>를 제시하였다. <표 3-9>

〈표 3-9〉 연도(조사차수)별 자녀 보육연령-출생연도 및 비용 지원 정책 대상 색인

유아학비 · 보육료 지원 연령 구분		보육 연령						
		0세		1세	2세	3세	4세	5세
가정양육수당 지원 월령 구분		12개월 미만	12~24 개월 미만	24~36 개월 미만	36~48 개월 미만	48~60 개월 미만	60~72 개월 미만	72개월 이상
기준연도	2010년(13차)	2010년생	2009년생	2008년생	2007년생	2006년생	2005년생	2004년생
	2011년(14차)	2011년생	2010년생	2009년생	2008년생	2007년생	2006년생	2005년생
	2012년(15차)	2012년생	2011년생	2010년생	2009년생	2008년생	2007년생	2006년생
	2013년(16차)	**2013년생**	**2012년생**	**2011년생**	**2010년생**	**2009년생**	**2008년생**	~~**2007년생**~~
	2014년(17차)	**2014년생**	**2013년생**	**2012년생**	**2011년생**	**2010년생**	**2009년생**	~~**2008년생**~~
	2015년(18차)	**2015년생**	**2014년생**	**2013년생**	**2012년생**	**2011년생**	**2010년생**	~~**2009년생**~~
	2016년(19차)	**2016년생**	**2015년생**	**2014년생**	**2013년생**	**2012년생**	**2011년생**	**2010년생**

주 : 1) 보육연령은 기준연도 1월 당시 자녀의 월령을 기준으로 하기 때문에, 0세에 두 해가 포함됨.
2) 각 셀마다의 연도는 자녀들의 출생연도를 의미하며, 음영 표시는 소득에 상관없이 전 계층에게 유아학비·보육료 지원이 이뤄진 시기의 출생연도, 음영 표시되지 않은 경우는 소득에 따라 지원 대상과 지원액이 구분되는 시기임.
3) 출생연도에 박스 처리가 된 경우는 양육수당이 지원된 시기로, 양육수당은 월령을 기준으로 지급되므로 기준연도에 일부 월이라도 양육수당을 받을 수 있는 경우는 박스 처리함. 영유아 연령과 소득에 따라 지원 대상이 제한적으로 적용된 시기는 박스 처리만 하였으며, 굵은 박스글씨로 처리된 경우에는 전 계층 양육수당이 지원됨.
4) 2013년(16차)~2016년(18차) 노동패널조사는 5세의 어린이집, 유치원 이용 행태를 질문하지 않은 차수로, 회고조사가 이뤄진 조사 대상을 사선으로 처리하여 표시함.

자료 : 최효미 외(2016), 「영유아 교육 · 보육비용 추정연구(Ⅳ)」, p.64. <표 Ⅲ-1-1>을 수정 보완함.

는 영유아 가구에 대한 유아학비·보육료 지원 및 가정양육수당 지원 정책이 각 연령별로 어떻게 적용되었는지를 보여주는 표로서, 박스 처리가 된 경우가 가정양육수당이 적용된 시기를 나타내며, 굵은 박스글씨로 처리된 경우가 전 계층에 대한 지원이 실시된 시기이다. 그리고 박스 처리가 된 시점과 연령이 전 계층에 유아학비·보육료 지원이 이뤄진 시기를 의미한다. 한편, 노동패널 회고조사 자료의 경우 커다란 정책 변화의 영향을 받는 연령이기 때문에, 결과 값에는 회고조사라는 특징뿐 아니라 정책 변화에 따른 효과가 섞여 있음에 주의할 필요가 있다.

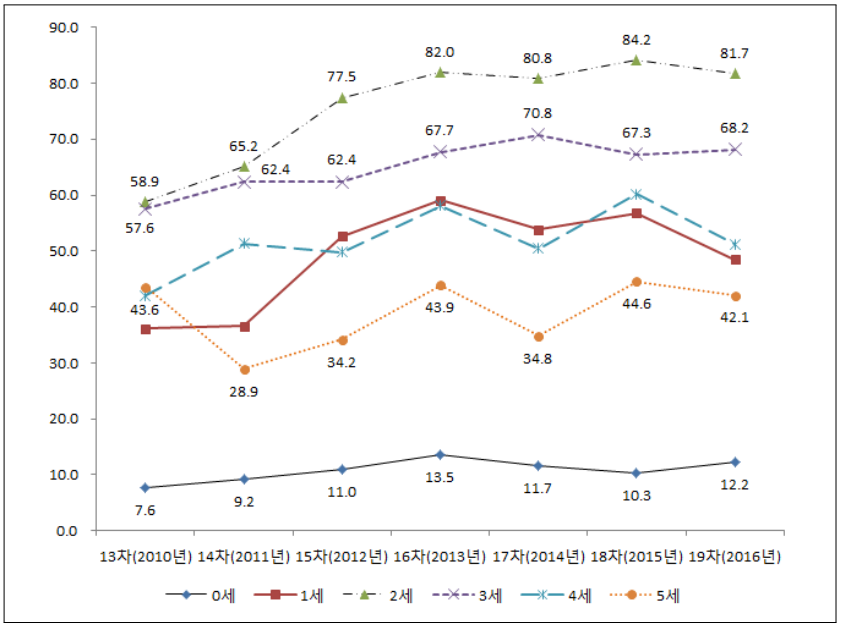
표에서 회고조사 자료는 사선 표시된 영역에 해당하는 경우로, 출생연도로는 2007년생~2009년생의 보육연령 5세(취학 전 1년)의 보육·교육 서비스 이용에 관한 사항이다. 표에서 보는 바와 같이 2007년생의 경우에는 4세까지는 정부로부터 유아학비·보육료 등의 지원을 받지 못한 코호트로, 회고조사가 이뤄진 5세(2013년, 16차조사)에 처음으로 유아학비·보육료 및 가정양육수당을 받을 수 있게 되었다. 반면, 17차(2014년)에 5세에 대한 회고조사가 이뤄진 2008년생은 전년도인 4세(2013년, 16차)에도 유아학비·보육료 지원 및 가정양육수당 지원을 받은 경우로, 전년도 자료는 회고조사가 아닌 본조사 당시에 응답이 완료된 경우이다.

16차~18차 사이의 5세아의 어린이집 이용 비중을 살펴보면, 16차 43.9%, 17차 34.8%, 18차 44.6%로 17차년도가 다른 연도에 비해 낮음을 알 수 있다. 그런데, 회고조사가 아닌 4세 아동의 어린이집 이용 비중 추세를 살펴보아도 17차년도에만 다소 낮아지는 경향성이 발견되며, 3세를 제외하고는 17차년도에는 모든 연령에서 전반적으로 어린이집 이용 비중이 줄어드는 양상을 보였다. 이러한 어린이집 이용 비중의 변화는 유아아동의 경우에는 어린이집의 대체재라 할 수 있는 유치원 이용 비중과의 관계 속에서 설명된다.

유아기 아동이 이용하는 유치원 이용 비중을 살펴보면, 5세를 기준으로 전체 분석 대상자 중 16차년도 49.8%, 17차년도 58.8%, 18차년도 46.7%가 유치원을 이용하고 있었다. 연령별로는 3세 이후 연령이 증가할수록 유치원 이용 비중이 증가하는 것이 일반적인 현상이며, 노동패널조사 자료 또한 이러한 경향성을 보이고 있다. 앞서 살펴본 바와 같이 17차(2014

〔그림 3-1〕 연도(조사차수)별, 연령별 어린이집 이용 아동 비중

(단위 : %)

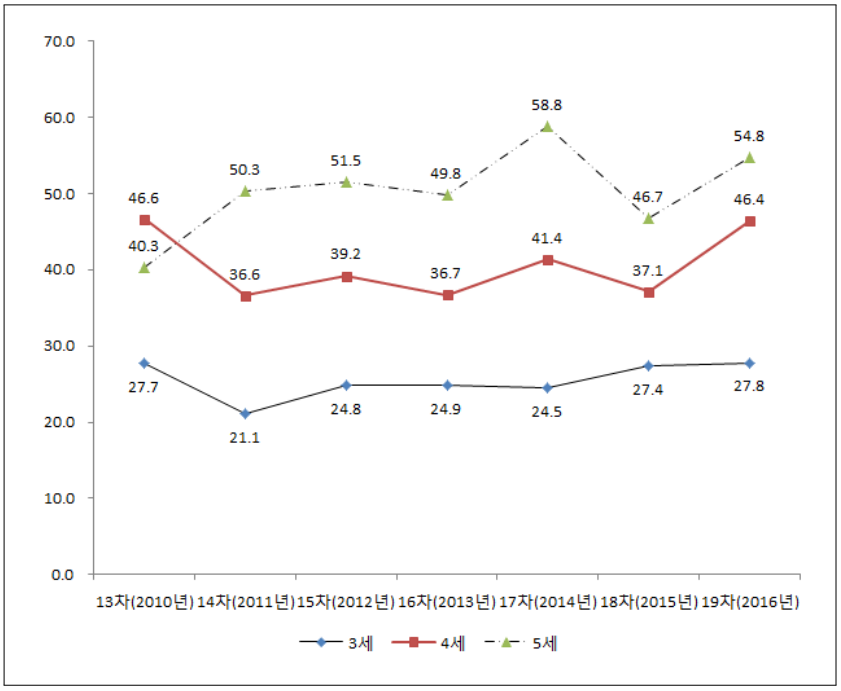


년)에는 4세와 5세 유아의 유치원 이용 비중이 다른 차수에 비해 크게 증가한 것으로 분석됐는데, 이는 어린이집 이용 비중이 적어진 것과 상관되어 있다. 즉 노동패널조사 자료의 경우 17차(2014년) 조사에서 어린이집보다 유치원 이용 아동의 비중이 크게 증가하는 양상을 보이는데, 이러한 경향은 5세에 국한된 현상이 아니라 다른 연령에서도 함께 관측되는 현상으로, 17차년도 5세 유아의 회고조사 자료의 문제라 보기는 어렵다.

연도별 5세의 개별돌봄서비스 이용 비중 변화를 살펴보면, 회고조사가 이뤄진 2013년(16차조사)에서 갑자기 개별돌봄서비스 이용자 비중이 크게 증가함을 알 수 있다. 이후 개별돌봄서비스 이용 아동의 비중이 다소 감소하기는 하지만 2012(15차조사) 이전에 응답자가 없거나 1명인 상황과 매우 상이함을 알 수 있다. 그런데, 이러한 현상은 비단 회고조사가 이뤄진 5세에 한해서 발생한 현상이라기보다는 당시에 해당 설문에 응답한 4세에서도 동일한 현상이 발견되었다. 따라서 이는 2013년(16차년도) 조사에서 설문의 구조가 크게 바뀌면서 발생한 현상으로 추측된다. 기존 설

〔그림 3-2〕 연도(조사차수)별, 연령별 유치원 이용 아동 비중

(단위 : %)



〈표 3-10〉 연도(조사차수)별, 연령별 개별돌봄서비스 이용 아동 비중

(단위 : 명, %)

		보육연령별 개별돌봄서비스 이용 비중					
		0세	1세	2세	3세	4세	5세
기준연도	2010년(13차)	6(2.3)	4(2.1)	1(0.5)	3(1.4)	0(0.0)	0(0.0)
	2011년(14차)	7(2.3)	2(1.1)	1(0.5)	0(0.0)	0(0.0)	0(0.0)
	2012년(15차)	6(2.0)	3(1.6)	2(1.1)	0(0.0)	1(0.5)	1(0.4)
	2013년(16차)	40(12.2)	34(16.4)	37(18.0)	24(11.9)	26(12.4)	38(18.5)
	2014년(17차)	30(13.0)	35(16.0)	22(10.3)	30(14.2)	24(12.1)	21(10.3)
	2015년(18차)	30(11.5)	24(13.6)	37(16.7)	24(11.5)	27(12.2)	23(11.8)
	2016년(19차)	23(10.8)	27(13.9)	28(16.0)	31(13.9)	31(14.8)	30(13.6)

주 : 음영부분은 보육교육비용이 응답되지 않은 경우로, 회고조사 자료를 기준으로 산출한 값임.

자료 : 한국노동연구원, 「한국노동패널조사」 원자료, 2010년(13차)~2016년(19차).

문에서도 동거 및 비동거 타인과 가족 친지는 동일 항목으로 조사되었으며 1~4번 응답 코드를 받았음에도 다수가 해당 항목에 응답을 누락하고 있었던 것으로 판단된다. 반면, 2013(16차년도) 조사 이후 아이돌봄서비스를 어린이집 혹은 유치원과 함께 응답하도록 함으로써, 이러한 차이가 발생한 것으로 보인다.

출생연도-보육연령별 어린이집+유치원 이용 비중을 살펴보면, 모든 연령에서 보육료·유아학비 지원이 시작된 해부터 어린이집+유치원 이용 비중이 급증하는 양상을 보였다. 2016년에 5세였던 2007년생의 경우 4세에 비해 5세 때 기관 이용이 다른 코호트에 비해 크게 증가하였는데,

〈표 3-11〉 출생연도-보육연령별 어린이집+유치원 이용 비중

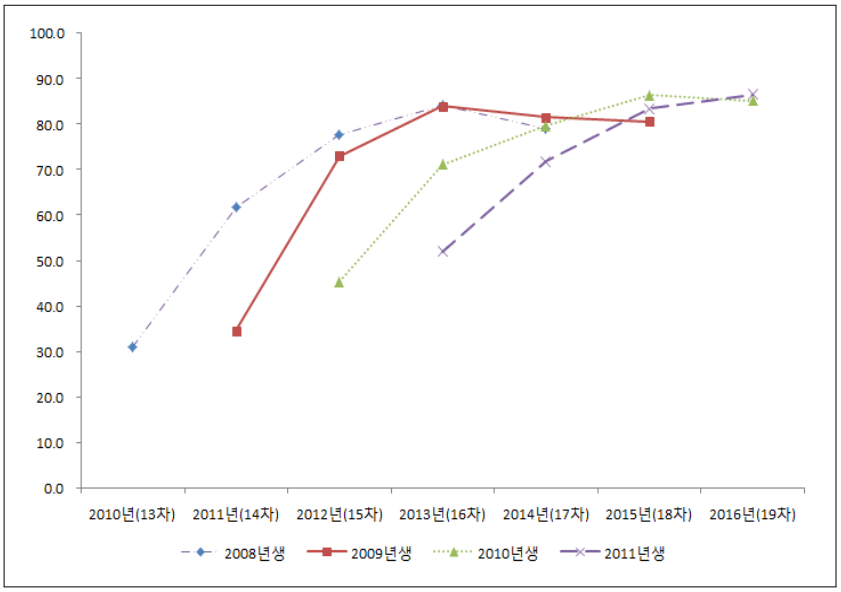
(단위 : 명, %)

		보육연령(월령)						
		0세		1세	2세	3세	4세	5세
		12개월 미만		12~24 개월 미만	24~36 개월 미만	36~48 개월 미만	48~60 개월 미만	60개월 이상
출생연도	2004년생	-	-	-	-	-	-	152 (84.0)
	2005년생	-	-	-	-	-	151 (83.4)	127 (70.2)
	2006년생	-	-	-	-	192 (78.4)	200 (81.6)	**194 (79.2)**
	2007년생	-	-	-	135 (60.0)	172 (76.4)	182 (80.9)	**200 (88.9)**
	2008년생	-	-	71 (31.1)	141 (61.8)	177 (77.6)	**192 (84.2)**	**180 (78.9)**
	2009년생	-	18 (8.5)	73 (34.6)	**154 (73.0)**	**177 (83.9)**	**172 (81.5)**	**170 (80.6)**
	2010년생	1 (0.4)	30 (12.7)	**107 (45.3)**	**168 (71.2)**	**188 (79.7)**	**204 (86.4)**	**201 (85.2)**
	2011년생	2 (0.9)	**31 (13.9)**	**116 (52.0)**	**160 (71.7)**	**186 (83.4)**	**193 (86.5)**	-
	2012년생	**4 (1.8)**	**41 (18.6)**	**114 (51.8)**	**178 (80.9)**	**196 (89.1)**	-	-
	2013년생	**5 (2.8)**	**26 (14.5)**	**96 (53.6)**	**140 (78.2)**	-	-	-
	2014년생	**0 (0.0)**	**19 (10.3)**	**87 (47.3)**	-	-	-	-
	2015년생	**3 (2.1)**	**26 (18.3)**	-	-	-	-	-

주 : 1) 회색 음영 처리된 부분은 2013년(16차년도) 조사의 응답 값이며, 굵은 글씨는 전 계층 유아학비·보육료 지원의 적용 대상이었던 경우를 나타냄.
2) 검정 바탕에 흰색 글씨로 표기된 2013년(16차년도) 이후 5세아의 자료는 회고 조사된 자료임.

자료 : 한국노동연구원, 「한국노동패널조사」 원자료, 2010년(13차)~2016년(19차).

〔그림 3-3〕 출생연도-연령별 어린이집 혹은 유치원 이용 아동 비중



주 : 1) 모든 코호트의 시작하는 연도의 보육연령은 1세부터이며, 2011년생은 2016년(19차) 조사 기준 5세에 도달하지 못하여 자료가 없음.
2) 모든 코호트에서 마지막 관측시점인 5세 때의 자료가 회고조사 자료임(2011년생은 5세 미관측).
자료 : 한국노동연구원, 「한국노동패널조사」 원자료, 2010년(13차)~2016년(19차).

이는 회고조사로 인한 오류라기보다는 당시 2007년생에게 보육료·유아학비 지원이 처음으로 시작됐기 때문으로 보인다. 반면, 전년도부터 보육료·유아학비를 지원받은 2008년생의 경우에는 4세 때 어린이집과 유치원 이용비중이 84.2%였으며, 5세 때 이용자 비중이 78.9%로, 다른 코호트와 유사한 패턴을 보였다.

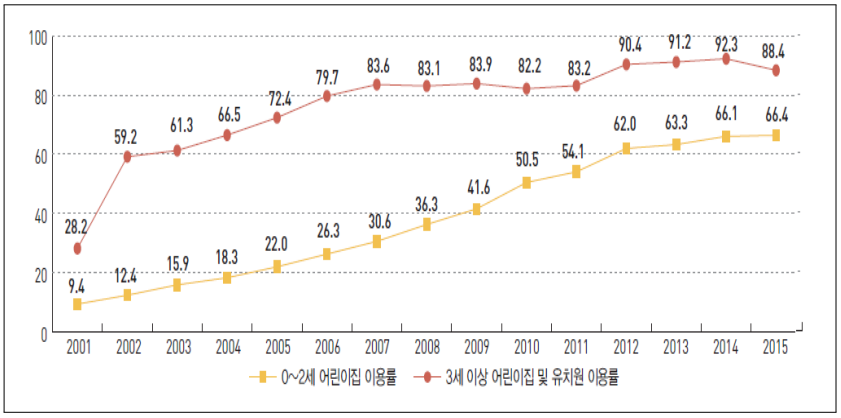
동일 연령인 5세를 기준으로 살펴보면 출생연도별로 어린이집+유치원 이용 비중이 80% 전후를 보이고 많이 이용하는 경우에 85% 정도였다. 예외라고 할 수 있는 수치는 2007년생이 5세인 2016년도로 이용자가 88.9%로 매우 높은 비중을 보이기는 하지만, 앞서 언급한 바와 같이 이는 정책 지원이 크게 확장된 시기로 인한 영향으로 사료되며 회고조사로 인한 회고오류라 보기는 어렵다. 지원 대상이 전 계층으로 확대가 된 2012년도와 2013년도를 전후하여, 기관 이용이 크게 증가하였다는 것은 잘 알

려진 사실이다(보건복지부・한국보건사회연구원, 2017: 219).

한편, 국내 어린이집 및 유치원 이용률 추이(보건복지부・한국보건사회연구원, 2017: 219)에 따르면, 2013년 0~2세 어린이집 이용률은 63.3%, 3세 이상 어린이집 및 유치원 이용률은 91.2%로, 노동패널과는 기관 이용 비중 산출 방식이 달라[4] 직접적 비교는 어렵지만, 전반적으로 노동패널 조사의 어린이집 및 유치원 이용 비중이 살짝 낮게 측정되는 양상을 보인다. 하지만, 이는 특정 연령이나 특정 해에만 국한된 현상이 아닌 전반적인 현상으로 노동패널 조사의 회고조사 자료만이 가진 독특한 특징은 아니었다.

〔그림 3-4〕 국내 어린이집 및 유치원 이용률(2001~2015)

(단위 : %)



자료 : 보건복지부・한국보건사회연구원(2017), 『통계로 보는 사회보장 2016』, p.219, [그림 3-8]을 재인용함.

2. 보육・교육비용 변화 추이

지금부터는 비용 변화 추이를 살펴보고자 한다. 노동패널조사는 2016년에 2013년 이후 조사되지 않은 5세의 어린이집 혹은 유치원 이용에 관한 설문을 회고 조사하였는데, 어린이집이나 유치원 이용여부 자체는 최

4) 보건복지부 자료는 연령을 영아(0~2세)와 유아(3~5세)로 구분하였는데 영아는 어린이집 이용률, 유아는 어린이집+유치원 이용률이다.

대 4년 정도 지난다 하더라도 회고가 어려운 응답항목이라 보기 어렵다. 반면, 어린이집이나 유치원 이용 비용은 4년 정도가 지나면 사실 회고가 어려울 수도 있기 때문에, 좀 더 주의가 요구되는 항목이라 할 수 있다.

어린이집을 이용하는 영유아 가구들 중에서 비용을 전혀 지불하지 않는다는 응답은 5세의 경우 대체로 5% 이하로 매우 낮은 비중을 보이는데, 회색으로 칠해진 회고자료의 경우에 최근 차수로 올수록 미지출 가구 비중은 더욱 적어지는 것으로 나타났으며, 이는 다른 연령에서도 공통적으로 보이는 경향이었다.

5세의 어린이집 이용 비용은 2012년에는 월평균 17만 5천 원이었으나 2013년 이후 회고자료는 30만 원 내외의 비용을 보였다. 2013년 5세에 대해서도 유아학비·보육료 지원이 이뤄진 시기이므로, 다소 의아한 결과라 할 수 있다. 다만, 5세뿐 아니라 다른 연령을 보더라도, 2013년에는 전반적으로 어린이집 이용 비용이 증가하는 기현상이 발생하고 있다. 다른 연령의 어린이집 이용 비용의 변화를 고려할 때, 회고조사 과정에서 비용조사에 오류가 있었다고 보기는 어려워 보인다. 이와 같은 현상이 발견되는 원인은 유치원 이용 비용까지를 살펴본 후에 다시 논의하고자 한다.

5세의 유치원 이용 비용도 2013년을 기점으로 급격히 증가하는 양상을 보인다. 유치원 이용 비용의 경우에도 어린이집과 유사하게 3세의 경우 2012년

〈표 3-12〉 연도(조사차수)별 보육·교육비용 미지출 가구

(단위 : 가구, %)

		보육연령별 미지출 가구					
		0세	1세	2세	3세	4세	5세
기준연도	2010년(13차)	231 (88.2)	101 (52.9)	50 (24.8)	19 (8.5)	8 (4.7)	9 (5.0)
	2011년(14차)	264 (86.0)	95 (51.9)	40 (20.4)	19 (9.3)	18 (7.9)	11 (6.8)
	2012년(15차)	261 (87.6)	85 (44.3)	30 (16.0)	17 (8.5)	11 (5.3)	9 (4.0)
	2013년(16차)	235 (79.7)	68 (35.1)	23 (11.7)	11 (5.7)	12 (5.9)	8 (4.2)
	2014년(17차)	161 (79.3)	67 (34.0)	25 (12.8)	10 (5.1)	12 (6.4)	7 (3.7)
	2015년(18차)	186 (81.2)	57 (35.0)	16 (7.9)	8 (4.1)	5 (2.3)	7 (3.8)
	2016년(19차)	146 (76.8)	74 (42.3)	20 (12.0)	7 (3.4)	3 (1.5)	6 (2.9)

주 : 음영부분은 보육교육비용이 응답되지 않은 경우로, 회고조사 자료를 기준으로 산출한 값임.

〈표 3-13〉 연도(조사차수)별 어린이집 월평균 이용비용(어린이집 이용자)

(단위 : 만 원, 명)

		보육연령별 어린이집 이용 비용					
		0세	1세	2세	3세	4세	5세
기준연도	2010년(13차)	25.9 (18)	22.6 (70)	23.1 (122)	24.5 (130)	24.4 (70)	20.7 (80)
	2011년(14차)	28.1 (30)	26.2 (67)	24.2 (130)	26.7 (131)	22.5 (117)	23.2 (46)
	2012년(15차)	11.0 (33)	12.2 (105)	14.5 (147)	21.9 (125)	20.6 (102)	17.5 (77)
	2013년(16차)	30.5 (44)	31.6 (115)	31.9 (162)	31.5 (130)	31.2 (116)	29.9 (86)
	2014년(17차)	34.8 (26)	33.2 (109)	30.8 (156)	31.8 (137)	30.7 (96)	31.3 (66)
	2015년(18차)	36.8 (22)	33.6 (94)	30.4 (171)	30.9 (132)	31.4 (123)	32.0 (83)
	2016년(19차)	37.1 (26)	35.9 (85)	32.9 (135)	34.0 (139)	33.4 (98)	34.0 (90)

주 : 1) 어린이집 이용자에 한해서 평균한 값이며(지불 금액이 0인 경우 포함), 괄호 안은 이용자수.
2) 음영부분은 회고조사 자료를 의미함.

〈표 3-14〉 연도(조사차수)별 유치원 월평균 이용비용(유치원 이용자)

(단위 : 만 원, 명)

		보육연령별 유치원 이용 비용		
		3세	4세	5세
기준연도	2010년(13차)	28.8 (62)	33.4 (80)	26.9 (73)
	2011년(14차)	34.8 (42)	34.0 (83)	30.5 (81)
	2012년(15차)	26.8 (52)	32.7 (80)	22.6 (118)
	2013년(16차)	37.8 (48)	36.9 (76)	40.2 (93)
	2014년(17차)	38.5 (50)	37.9 (77)	38.8 (114)
	2015년(18차)	35.4 (54)	38.8 (81)	39.8 (87)
	2016년(19차)	45.2 (57)	39.4 (95)	41.6 (111)

주 : 유치원 이용자에 한해서 평균한 값이며(지불 금액이 0인 경우 포함), 괄호 안은 이용자수.

대비 2013년에 급격히 상승하는 현상을 보이고 있다. 다만, 4세의 경우에는 큰 증가를 보이지 않았다. 2012년에는 3세와 4세에 대해서만 정부 지원이 이뤄지지 않다가 2013년도에 정책이 확장된 연령인데도 정책이 확장된 해인 2013년에도 유아학비가 줄지 않았다.

이러한 추세를 보다 자세히 살펴보기 위해 출생연도-보육연령별로 어린이집 이용비용(표 3-15 참조)과 유치원 이용비용(표 3-16 참조)을 구해 보았다. 먼저, 어린이집 이용비용의 경우에는 앞서 언급한 바와 같이 2006년생(2012년 조사)과 2007년생(2013년 조사) 사이에 큰 상승이 발견되었다. 그런데, 같은 출생연도 아이들의 연령 증가에 따른 변화를 살펴보면, 2007년생의 경우에는 4세 때 오히려 비용이 살짝 감소하였다가 5세에 다시 증가하는 양상을 보이는 반면, 2008년생의 경우 3세에서 4세로 가는 시점에 비용 증가가 컸고 5세 때는 크지 않은 양상을 보였다. 2009년생은 그 시기가 더 빨라져 2세에서 3세로 가는 과정에서 비용 증가가 컸으며, 2010년생 1세에서 2세 사이가 비용 증가가 큰 양상을 보였다. 이를 간략히 정리하자면, 출생연도를 막론하고 2012년에 낮아진 어린이집

〈표 3-15〉 출생연도-보육연령별 어린이집 이용 비용

(단위 : 만 원, 명)

		보육연령						
		0세		1세	2세	3세	4세	5세
출생연도	2004년생	-	-	-	-	-	-	20.7 (79)
	2005년생	-	-	-	-	-	24.4 (71)	23.2 (46)
	2006년생	-	-	-	-	24.5 (130)	22.5 (117)	17.5 (76)
	2007년생	-	-	-	23.1 (122)	26.7 (130)	20.6 (102)	29.9 (86)
	2008년생	-	-	22.6 (70)	24.2 (130)	21.9 (125)	31.2 (116)	31.3 (66)
	2009년생	-	25.6 (17)	26.2 (67)	14.5 (147)	31.5 (129)	30.7 (95)	32.0 (83)
	2010년생	31.0 (1)	27.6 (28)	12.2 (105)	31.9 (162)	31.8 (138)	31.4 (123)	34.0 (90)
	2011년생	34.5 (2)	11.1 (29)	31.6 (115)	30.8 (156)	30.9 (132)	33.4 (98)	-
	2012년생	10.0 (4)	29.8 (39)	33.2 (110)	30.4 (172)	34.0 (139)	-	-
	2013년생	35.6 (5)	34.8 (26)	33.6 (95)	32.9 (135)	-	-	-
	2014년생	-(0)	37.3 (19)	35.9 (85)	-	-	-	-
	2015년생	33.7 (3)	37.1 (26)	-				

주 : 1) 회색 음영 처리된 부분은 2013년(16차년도) 조사의 응답 값이며, 굵은 글씨는 전 계층 유아학비·보육료 지원의 적용 대상이었던 경우를 나타냄.
2) 검정 바탕에 흰색 글씨로 표기된 2013년(16차년도) 이후 5세아의 자료는 회고 조사된 자료임.

자료 : 한국노동연구원, 「한국노동패널조사」 원자료, 2010년(13차)~2016년(19차).

〈표 3-16〉 출생연도-보육연령별 교육·보육서비스 유치원 이용 비용

(단위 : 만 원, 명)

		보육연령		
		3세	4세	5세
출생연도	2004년생	-	-	26.9 (73)
	2005년생	-	33.4 (80)	30.5 (81)
	2006년생	28.8 (62)	34.0 (83)	**22.6 (118)**
	2007년생	34.8 (42)	32.7 (80)	**40.2 (93)**
	2008년생	26.8 (52)	**36.9 (76)**	**38.8 (114)**
	2009년생	**37.9 (48)**	**37.9 (77)**	**39.8 (87)**
	2010년생	**38.5 (50)**	**38.8 (81)**	**41.6 (111)**
	2011년생	**35.4 (54)**	**39.4 (95)**	-
	2012년생	**45.2 (57)**	-	-

주 : 1) 회색 음영 처리된 부분은 2013년(16차년도) 조사의 응답 값이며, 굵은 글씨는 전 계층 유아학비·보육료 지원의 적용 대상이었던 경우를 나타냄.
2) 검정 바탕에 흰색 글씨로 표기된 2013년(16차년도) 이후 5세아의 자료는 회고 조사된 자료임.

자료 : 한국노동연구원, 「한국노동패널조사」 원자료, 2010년(13차)~2016년(19차).

혹은 유치원 이용 비용이 2013년에는 높아진다고 정리할 수 있다. 2012년은 3세와 4세를 제외하고는 전 소득 구간에 유아학비・보육료 지원이 실시된 시기로, 3세와 4세의 비용도 동반 하락하는 현상이 발생하였다. 반면, 2013년은 3세와 4세를 포함하여 전 소득 구간에 대해 정부 지원이 이뤄졌음에도 어린이집 이용비용이 전반적으로 상승하는 현상이 발견된다.

이상의 결과에서 회고조사 자료의 경우 이전 자료에 비해 비용이 크게 증가하는 현상이 발견되지만, 이는 다른 연령에서도 동일하게 발견되는 현상으로 회고조사 자료의 오류라기보다 정책 지원의 역작용이 반영된 결과로 보인다. 즉 이처럼 유아학비・보육료 지원이 크게 확장되었음에도 불구하고, 영유아 가구의 보육・교육비용이 감소되지 않은 현상은 일종의 정책 효과가 다른 요인에 의해 상쇄되는 현상을 반증하는 결과인데, 유치원의 경우에는 유아학비 지원(누리과정 지원금) 이외에도 별도의 수업료를 받을 수 있도록 되어 있어, 제도 시행 첫해를 제외하고 수업료가 크게 줄어들지 않은 현상에 기인한다. 한편, 어린이집의 경우에는 유치원

에 비해 보육료 수납 원칙이 상당히 엄격하게 적용되고 있는 기관이긴 하나, 어린이집이 경우에도 보육료 지원금만으로 원의 운영이 불가능한 상황을 고려하여 각종 필요경비 및 특별활동 비용 등에 대해 부모 부담이 가능하도록 되어 있어, 이로 인한 비용 절감 효과가 감소한 결과로 보인다[5]. 이처럼 어린이집이나 유치원에 막대한 재정이 투입됨에도 불구하고, 유아학비・보육료가 기관에 대한 지원이 아닌 가구에 대한 지원금의 성격을 갖고 있는 데다 이용자 급증에 따라 어린이집과 유치원이 수익자 부담금을 부과하게 됨으로써 정책 지원금의 효과가 상쇄되고 있다는 것은 이미 잘 알려진 사실이기도 하다(최성은, 2016; 서수경, 2012 등).

요컨대, 노동패널조사 회고조사 자료에서 나타나는 이용자 비중이 크게 증가하거나 이용비용에서 보이는 특이한 변동은 응답자의 회고 오류로 인한 변동이라기보다는 정책 변동에 따른 소비자 행태 변화를 반영한 것으로 이해될 수 있다. 이는 유아학비・보육료 지원 이후에도 어린이집이나 유치원 이용 비용이 크게 감소하지 않았을 뿐 아니라, 어린이집 혹은 유치원의 이용이 급증하고 비용이 증가했다는 선행연구(김은정・이혜숙, 2016; 민연경・장한나, 2015; 안현미・박애리, 2013 등)의 결과와 맥을 같이한다.

3. 보육・교육비용 결정요인

지금까지 노동패널조사 회고조사 자료의 특징에 대해 살펴보았는데, 실제로 회귀분석 등을 실시하는 과정에서 회고조사 자료가 특이한 현상을 보이는지를 알아보기 위해 본 연구에서는 간단하게 만 5세의 어린이집과 유치원 이용비용을 종속변수로 하는 회귀분석을 실시해 보았다. 분석은 두 가지 경우로 나누어 실시되었는데, 15차년도 이전과 16차년도 이후 만 5세의 어린이집과 유치원 이용비용에 영향을 미치는 요인을 각각 분석해 보았으며, 동일 시점에 연령별로도 차이를 보일 수 있으므로 만 5세

5) 정부 지원금의 증가로 인해 보육서비스 이용 필요성이 낮은 학부모들까지 어린이집 혹은 유치원을 이용하면서, 이용자가 급증하였을 뿐 아니라 장시간 기관 보육이 확산되면서 어린이집 운영비 자체가 증가했다는 것이 현장의 의견이다.

와 가장 유사한 양태를 보이는 연령인 만 4세의 어린이집과 유치원 이용 비용에 영향을 미치는 요인을 비교 분석해 보았다. 이때, 보육·교육비용에는 모의 취업 상태 등 특성이 주로 영향을 미치는 것으로 알려져 있으나, 본고에서는 관측치를 최대한 많이 살리기 위해 모의 특성은 결정 요인 분석에 포함시키지 않고, 주요 아동 특성과 가구 특성을 중심으로 간략한 분석을 실시하였다는 점을 먼저 밝혀둔다.

먼저, 만 5세의 영유아 교육·보육비용 지출 결정 요인을 살펴보면, 15차년도 이전 자료는 아동의 출생순위가 둘째인 경우에는 첫째 자녀에 비해 기관 이용 비용 지출이 유의미하게 감소했으며, 자녀수가 많을수록 지출 비용이 감소하는 것으로 분석되었다. 반면, 16차년도 이후 회고조사 자료에서는 이러한 출생순위에 따르는 유의미한 효과가 발견되지 않았다. 한편, 이용 기준이 어린이집인 경우에 비해 유치원에 다니는 경우에 유의미하게 비용 지출이 증가함을 알 수 있으며, 가구소득에 따라서 소득이 높을수록 비용 지출이 높은 경향을 보인다. 단, 15차 이전 조사에서는 300만 원 이상의 모든 소득 구간에서 유의한 효과가 발견된 반면, 16차년도 이후 회고조사 자료에서는 500만 원 이상의 소득 구간에서만 유의미한 차이가 발견되어, 전반적으로 가구소득이 고소득인 경우에는 유의한 차이를 보이고 있음을 알 수 있다.

이러한 차이가 회고조사 자료에서만 발견되는 현상인지, 16차년도 이후에 발견되는 일반적 현상인지를 검토해 보기 위해 만 4세에 대한 분석을 실시해 본 결과, 16차년도 이후 가구소득의 효과가 고소득 구간에서만 나타나는 현상은 회고조사 자료가 아닌 만 4세에서도 동일하게 발견되는 현상이었다. 단, 16차년도 이후 만 4세에 대한 분석 결과에서는 아동의 출생 순위가 높아질수록 지출 비용이 감소하는 현상이 포착되었다는 점이 만 5세와의 차이점이다. 이때 아동의 출생순위가 지출 비용에 영향을 미친다는 결과는 15차년도 이전 만 4세에 대한 회귀분석 결과에서도 동일하게 발견되는 현상이며, 15차년도 이전 만 5세에 대한 분석 결과에서도 둘째 자녀인 경우에는 유의한 효과가 발견되기는 했지만, 자녀의 연령이 어릴수록 출생순위가 높은 아이들의 기관 이용 빈도가 낮아지는 현상이 반영된 결과로 보인다.

즉 회귀분석 결과, 16차년도 이후 만 5세의 교육·보육비용 지출에 영향을 미친 요인이 15차 이전 조사 및 만 4세에는 약간 상이한 결과를 보이고 있기는 하나, 전반적인 경향상에서 크게 오류가 있는 것으로 보이지는 않는다. 16차년도 이후 가구소득의 효과가 고소득 가구에 국한된 현상은 회고조사 자료에서만 발견되는 현상이 아니라 만 4세에서도 동일하게 발견되는 현상으로, 16차년도(2013년도)에 정책 지원금의 확대로 인한 효과가 반영된 결과로 사료된다.

〈표 3-17〉 영유아 교육·보육비용 지출 결정요인

		만 5세		만 4세	
		13차~15차	16차~18차	13차~15차	16차~18차
아동 성별 (기준=남아)		1.338	-0.197	-0.521	1.484
아동 출생순위 (기준=첫째 자녀)	둘째 자녀	-4.768**	-0.118	-3.588*	-2.399†
	셋째 자녀 이상	-1.986	-0.525	2.205†	-2.867**
총 자녀 수(연속변수)		-3.085*	-0.176	4.965**	-1.901
이용 기관(기준 = 어린이집)		5.600***	7.961***	8.890***	6.566***
가구소득 (기준= 199만 원 이하)	200~299만 원	0.540	2.288	2.927	1.396
	300~399만 원	8.033**	2.238	4.880†	0.525
	400~499만 원	7.161*	1.920	8.178**	3.320
	500~599만 원	11.302**	4.446*	16.061***	2.539
	600만 원 이상	18.429***	7.135***	26.481***	10.007***
조사 차수	14차	1.620		0.413	
	15차	3.815*		-2.758	
	17차		-0.241		-0.320
	18차		0.437		0.287
상수		22.952***	28.727***	28.098***	32.871***
N		491	561	553	595

주 : †p < .1, * p < .05, ** p < .01, *** p < .001

제5절 사교육비 부담과 관련된 문항에 대한 검토

한편, 16차년도 이후 설문에 포함된 사교육비 부담과 관련된 문항의 경우에는 가구 단위 지출에 대한 사교육비 부담을 질문한 문항으로, 회고조사 자료로 인한 결과의 변동이 있었는지 검토가 필요한 상황이다. 즉 16차년도 조사 당시 영유아가 있는 가구 부모들은 가구에서 모든 영유아에게 지출하는 사교육비가 부담스러운지를 질문하였는데, 조사 당시 만 5세 아동들은 해당 설문에 응답되지 않았기 때문에, 당시 가구 내 영유아가 만 5세아 한 명인 경우에는 해당 문항에 응답 값이 없는 상황이다. 반대로, 조사 당시 가구 내 만 5세아를 포함하여 더 어린 연령의 영유아가 있었던 경우라면 해당 문항에 응답을 했을 테지만, 앞서 문항에서 만 5세아의 보육·교육서비스 이용 관련 문항을 응답하지 않았기 때문에 해당 문항은 가구에서 지출하는 비용 전체에 대한 부담을 질문하고 있음에도 만 5세아에게 지출되는 비용을 감안하지 않고 응답했을 가능성을 완전히 배제하기 어렵다.

앞서 언급한 전자의 경우, 즉 조사 당시 가구에 영유아가 만 5세아밖에 없어서 해당 문항에 응답이 누락된 경우에는 회고조사 자료의 값으로 누락된 응답 값을 채워주기만 하면 되기 때문에 비교적 간단하게 문제가 해결된다. 하지만, 후자의 경우, 즉 영유아 가구가 해당 비용을 가구 단위로 인지하지 못한 채 사교육비 부담을 응답한 경우에는 회고조사 자료와 큰 차이를 보이고 있는지를 검토해 볼 필요가 있다.

노동패널조사에서 16차년도 이후 추가된 사교육비 부담 관련 문항은 사교육비의 경제적 부담 정도에 대한 문항과 정부 지원의 사교육비 부담 경감 정도에 관한 문항으로 2개이다. 조사 당시 사교육비의 경제적 부담에 대한 응답 값과 회고조사 당시의 응답 값이 차이를 보이는 경우는 16차년도 12가구였으며, 18차년도는 5가구였다. 한편, 정부 지원의 사교육비 부담 경감 정도에 관한 문항에서 조사 시점에 따라 차이를 보이는 문항은 16차년도 5가구, 18차년도 7가구였다. 각 차수별로 회고조사 자료가

〈표 3-18〉 영유아 자녀 사교육비 관련 문항 응답이 차이를 보이는 경우(원자료 vs 회고자료)

(단위 : 가구)

	사교육비 경제적 부담 정도	정부 지원의 사교육비 부담 경감 정도
16차년도	12	5
17차년도	0	0
18차년도	5	7

자료 : 한국노동연구원, 「한국노동패널조사」 원자료, 2010년(13차)~2016년(19차).

약 200 내외인 점을 감안할 때 해당 문항에 조사 시점에 따라 다른 응답을 한 가구수가 많지는 않은 상황이다. 또한, 구체적인 응답 내용을 가구별로 비교해 보면, 응답 값이 큰 차이를 보이지는 않았다. 응답 값이 차이를 보이는 경우에 대체로 조사 당시에는 '매우 부담스럽다(5)'고 응답했으나, 회고조사에는 '부담됨(4)' 정도로 응답하는 경우와 같이 척도 점수가 큰 차이를 보이지 않는 경우가 대부분이며, '매우 부담스럽다(5)'고 응답했다가 '전혀 부담되지 않는다(1)' 등과 같이 크게 변동한 경우는 찾아보기 어려웠다.

사교육비 부담 정도와 정부지원의 부담 경감 도움 정도와 같은 문항은 주관적 판단을 묻는 문항이므로, 객관적으로 두 자료 중 어느 쪽에 더 정확한 조사라고 말하기 어려운 한계가 있다. 다만, 조사 당시 가구 내 모든 영유아의 교육·보육서비스 이용을 묻지 않았기 때문에 이후 가구 단위 조사에서 누락된 자녀에게 소요되는 비용을 감안하지 않고 응답하였다는 문제는 회고조사 자료에도 동일하게 존재하는 문제이다. 즉 회고조사 자료는 조사가 누락된 만 5세 아동에 대해서만 추가적인 질문을 실시한 것으로, 이 경우가 조사 당해연도에 비해 오히려 누락된 자녀에 대한 교육·보육서비스만을 생각하며 가구 단위 조사에 응답했을 가능성이 높은 것으로 보인다.

그러므로, 전반적으로 조사 당시 원자료의 해당 문항에 응답 값이 있는 경우는 그것을 회고조사 자료에 비해 우선하고, 당시 영유아 자녀가 만 5세 자녀 한 명만 있어서 해당 문항 값이 아예 조사되지 않은 경우(응답 값이 미싱인 경우)에는 회고조사 자료의 값을 보조적으로 활용하여 채워

넣는 것이 보다 타당하다고 사료된다.

제6절 소 결

이상의 논의를 간략히 요약하고, 이를 통한 시사점을 도출하면 다음과 같다. 먼저, 노동패널조사는 16차년도(2013년) 조사에서 보육연령 5세(취학 전 1년)의 보육·교육서비스 이용 및 비용에 관한 조사가 누락되어, 19차년도(2016년) 조사 당시 회고조사를 통해 자료를 보완하였다. 하지만, CAPI상에서의 로직 오류로 발생한 응답 대상자의 조사 누락은 단순한 기술적 문제라기보다는 16차년도 당시 설문지 수정 과정에 따르는 변화와 정책 환경의 변화 등을 포함하는 복합적 문제라 볼 수 있다.

회고조사 자료와 기존 자료를 통해 살펴본 보육연령 5세의 보육·교육서비스 이용 및 비용 변화 추세를 살펴보면, 실제 16차년도(2013년)에는 상당한 정도의 변화가 포착되었다. 하지만, 이러한 변화는 회고로 인한 응답 오류라기보다는 상당부분이 정책 환경 변화를 반영하고 있는 것으로 추측된다. 응답 누락이 발생한 2013년도는 영유아기 전 연령에 대해 소득수준과 무관하게 유아학비·보육료 지원 및 가정양육수당 지원이 시작된 해로서, 전반적으로 교육·보육서비스 이용 행태와 비용 변화가 급격한 해에 속한다. 이러한 변화는 만 0~2세와 5세에 대해 영유아 교육·보육비용이 보편적으로 지원된 2012년부터 가속되는 현상으로, 2013년도에 새롭게 정책 수혜를 받게 된 만 3~4세에서도 동일하게 발견되는 현상이다. 즉 노동패널 16차년도(2013년) 회고조사 자료는 이전 연도와의 추세를 비교해 볼 때 어린이집 혹은 유치원 이용이 급격히 증가하는 등의 변화가 관측되지만, 이러한 현상은 당해연도 조사가 완료된 만 3~4세에게서도 동일하게 관측되는 현상일 뿐 아니라 정책 변화로 인해 이용 행태가 크게 변동한 전년도(2012년) 자료와 유사한 패턴을 보이는 현상이라 볼 수 있다. 또한, 17차년도(2014년) 이후 서비스 이용 비율과 비용이 살짝 감소하였다가 다시 증가하는 양상 또한 선행 연구들의 결과를 고려할

때, 응답자의 회고에 따른 오차라기보다는 정책 환경의 변화에 따른 것으로 크게 문제 되는 수준은 아니라고 할 수 있다. 즉 전반적으로 16차년도~18차년도에 걸친 노동패널 보육연령 5세에 대한 회고조사 자료는 조사과정과 결과에 있어서는 큰 문제가 없다고 판단된다.

또한, 16차년도(2013년) 조사 당시 응답 가구가 회고조사에 모두 응답한 것은 아니라서 응답 누락으로 인한 변동이 일부 포함되었을 가능성이 있으나, 출생연도를 기준으로 한 코호트 분석에서도 특이사항이 발견되지 않는 등 이로 인한 문제도 크지 않을 것으로 사료된다.

하지만, 16차년도 이후 노동패널 보육·교육서비스 관련 자료는 회고조사로 인한 문제보다는 설문구조의 변경에 따른 문제가 오히려 큰 것으로 보이며, 향후에는 이에 대한 개선이 반드시 요구된다. 16차년도 이후 노동패널조사는 영유아의 교육·보육서비스 중에서 어린이집과 유치원, 개별돌봄서비스와 같은 일부 서비스에 대해서만 조사를 실시하였다. 이 과정에서 종일제 개별돌봄서비스의 경우에는 이용 비중이 크게 변동하는 등의 문제가 발견되었는데, 이는 회고조사의 오류라기보다는 설문구조 변경에 따라 응답자들의 응답 내용이 바뀐 것으로 판단된다. 응답자들이 이전에 비해 개별돌봄서비스에 보다 많이 응답하게 된 원인에 대해서는 노동패널조사 자료만을 통해서는 판단하기 어렵지만, 16차년도 이후 해당 서비스 이용 비중이 크게 증가했다기보다는 15차년도 이전에 20개 이상의 항목 중에서 보육·교육서비스를 복수 응답하도록 한 기존의 설문구조로 인해 응답 누락이 있었다고 보는 것이 보다 타당해 보인다. 즉 16차년도 이후 설문의 구조가 응답자의 응답 편의를 제고하였다고 볼 수 있다. 따라서 설문의 구성은 16차년도 이후와 같이 설문을 구분하여 질문하는 것이 보다 응답 편의를 제고할 수 있을 것으로 사료된다.

다만, 16차년도 이후 설문이 보다 응답에 용이한 구조를 가졌다 할지라도, 이로 인해 15차년도 이전과 16차년도 이후 개별돌봄서비스 이용에 관한 자료의 연속성은 확보하기 어려운 상황이다. 또한, 16차년도 이후로는 영유아의 교육·보육서비스 이용에 있어 사교육(학습지, 학원, 문화센터 등)에 대한 설문이 삭제되어 조사가 이루어지지 않았기 때문에 15차년도 이전과는 자료가 큰 차이를 보인다. 그럼에도 불구하고, 시계열 연속성과

변수의 간소화를 위해 16차년도 이후 자료도 기존 자료와 같은 형태로 조사된 자료를 재가공하여 제공하고 있기 때문에 오히려 자료 사용에 있어 오해와 혼선을 가중할 위험이 높은 것으로 보인다. 쉽게 말해, 서비스 종류를 구분하지 않고 영유아 아동에 대한 교육·보육비를 산출하고자 하는 연구자가 있다면, 15차년도 이전 자료까지는 사교육비가 포함되는 대신 개별돌봄서비스 이용 비용은 과소하게 책정되었을 가능성이 높은 반면 16차년도 이후 자료는 사교육비가 아예 반영되지 않는 대신 개별돌봄서비스 비용은 크게 잡혀 있는 상황이 발생한다는 것이다. 그럼에도 불구하고, 현행 노동패널 통합설문지 및 코드북, 자료 등에는 이러한 상황이 자세하게 설명되어 있지 않고, 모든 차수에서 동일한 변수로 자료가 처리되어 있기 때문에 자료의 오해석이 발생할 위험이 높다.

요컨대, 영유아 교육·보육 관련 분석에서 15차년도 이전과 16차년도 이후 자료에서 연속성을 가지고 분석할 수 있는 자료는 어린이집과 유치원 이용 비율 및 비용에 국한된다. 그러므로, 노동패널 자료 사용에 있어서의 오해를 방지하고 이용 편의를 제고하기 위해서는 기존의 데이터 구조보다는 어린이집 이용자, 유치원 이용자 등으로 구분하여 자료를 제공하는 방안을 검토해봄 직하다. 이때, 15차년도 이전 자료의 경우에는 그 외의 서비스가 조사되었기 때문에 이를 전부 서비스별로 전환하려면 변수가 지나치게 증가하게 되는 등의 문제가 발생한다. 따라서 보다 세분된 서비스 종류에 대한 분석을 희망하는 연구자들을 위해서 기존 데이터 구조를 가진 자료를 패치파일로 제공하고, 이후로는 16차년도 설문 변경의 내용을 보다 잘 반영할 수 있는 데이터 구조를 결정하여 원자료와 함께 제공하는 방안에 대한 전향적 검토가 요구된다. 단, 노동패널조사는 20차년도 조사 이후에도 정책 변화 환경을 반영하기 위해 영유아 교육·보육 비용에 대한 설문구조의 변경 등을 고려하고 있으므로, 데이터 구조의 변화는 향후에 있을 설문의 변화를 충분히 반영하여 차후에 한꺼번에 진행되는 것이 바람직할 것으로 사료된다.

따라서 조사 방식 등의 전면적 검토가 실시되기 이전까지는 원자료는 기존 방식과 같은 방식으로 통합하여 데이터를 제공하되, 자료 특성에 대한 충분한 설명을 첨부하여 오해석을 낳지 않도록 해야 할 것이다. 또한,

사교육비의 경제적 부담 정도 및 정부 지원의 사교육비 경감 정도 문항은 16차년도 이후 추가 신설된 문항으로, 조사 당시 가구 내 만 5세 유아 이외에 다른 영유아가 없어 응답이 누락된 경우에는 회고조사에서 응답된 내용을 보완적으로 추가하여 자료를 가공할 필요가 있다. 그리고 기존 방식으로 원자료를 재구성하여 통합 데이터를 제공하는 경우라 할지라도 회고조사로 진행된 자료에 대해서는 회고조사 자료임을 명확하게 알 수 있도록 '회고조사 여부' 변수를 신설하여 함께 제공함으로써, 회고조사 자료를 사용하기를 꺼려는 유저들은 회고조사 자료를 지우고 연구를 진행할 수 있도록 지원하는 것이 바람직하다고 사료된다.

제 4 장
소득세 추정 모형(Personal income tax calculator)의 필요성과 활용 방안

제1절 서 론

일반적으로 소득세추정모형(Personal income tax calculator)은 미시 횡단면 혹은 패널자료를 주로 활용하여 현행 소득세제하에서 개인·가구들이 부담하는 납세액을 산출(calculate)하는 모형을 의미한다. 나아가 현실에서 논의되고 있는 세제 개혁안들에 따른 개인·가구들의 납세액과 그 변화를 시뮬레이션 하기도 한다. 이를 통해 세제 개편이 세수에 미치는 영향은 물론, 소득세제의 재분배 기능에 미치는 효과도 분석할 수 있다. 특히, 세제 개혁으로 인한 이득/손실이 상호 이질적인 개인들 사이에서 어떻게 나타날 것인지를 구체적으로 파악할 수 있다는 점에서 정책적으로 매우 유용하다.

다른 한편, 소득세추정모형을 활용할 때 세제개혁이 개인·가구의 한계유효세율(marginal effective tax rate)에 미치는 영향을 추정할 수 있는데, 이것은 정책적 차원은 물론, 학문적 차원에서도 매우 중요하다. 왜냐하면 세제개혁으로 인한 한계유효세율의 변화는 개인의 근로 및 저축 등과 같은 경제적 유인에 영향을 미치는 주요 요인이기 때문이다. 다시 말하면, 소득세추정모형의 개발과 활용은 과세(taxation)가 개인·가구의

경제·사회적 행위에 미치는 영향에 대한 다양한 연구를 촉진·발전시킬 수 있다는 점에서 주목할 만하다.

이러한 맥락에서 많은 국가와 연구기관에서 소득세추정모형(tax calculator) 혹은 조세/복지급여 모형(tax/benefit model)을 개발·활용하고 있다. 하지만 아직까지 우리나라에서는 이와 관련된 모형의 개발과 활용은 찾아보기 힘든 것은 물론, 해당 모형에 대한 관심조차 크지 않은 실정이다. 이에 본 연구는 소득세추정모형의 필요성과 의의 등에 대해서 살펴보는 한편, 지난 2년 전부터 노동패널 자료를 활용하여 개발하고 있는 소득세추정모형의 주요 결과를 제시·검토하고자 한다.

좀 더 구체적인 글의 구성은 다음과 같다. 다음 제2절에서는 소득세추정모형의 개발과 활용의 필요성과 이와 관련된 국내·외 사례를 살펴보고 시사점을 도출하고자 한다. 다음으로 제3절에서는 지난 2년 동안 수행된 2007~2014년까지의 소득세 추정결과를 2015년까지 연장한 결과를 검토하고, 나아가 모형의 유효성(validation)을 검토하고자 한다. 마지막으로 제4절에서는 본 연구에서의 소득세 추정의 한계와 향후 발전방향에 대해 논의하면서 글을 마치고자 한다.

제2절 소득세 추정모형의 필요성과 국내·외 사례

1. 소득세 추정모형의 의의와 필요성

경제학에서 조세는 기본적으로 시장의 효율성을 침해한다고 간주된다. 왜냐하면 특정 경제활동에 대한 어떠한 과세도 해당 조세와 관련된 시장의 가격기능을 왜곡시킴으로써 자원배분의 비효율성을 발생시킬 수 있기 때문이다. 또한 개인의 납부능력이 아닌, 경제적 중요성에 대한 정책결정자의 판단에 따라 특정 집단의 세부담을 축소/증가시킴으로써 형평성은 물론, 시장의 효율성을 침해할 수 있다. 예를 들어, 고소득자/대기업이 더 많은 혜택을 받는다거나(수직적 형평성 침해), 경제적 능력이 유사한 개

인들이지만 실제로 직면하는 유효세율이 다를 수 있다(수평적 형평성 침해). 이러한 형평성 침해가 시장의 신호기능을 왜곡시킬 수 있고 그 결과 효율성을 축소시킬 수 있기 때문이다.

그럼에도 불구하고 다양한 정책 목적들을 달성하기 위해 많은 국가에서 조세제도를 활용하는데 그 이유는 다음과 같다.

첫째, 사회복지(social welfare)는 경제적 효율성을 넘어선 사회적 가치를 포괄한다는 점에서, 과세에 의한 효율성 침해가 존재한다고 할지라도 형평성 증진을 통해 사회복지가 증가될 수 있다.

둘째, 외부성을 갖는 경제활동 촉진을 통해 사회복지를 증가시킬 수 있다. 예를 들어, 특정 지역 혹은 집단의 지원을 통한 집적(agglomeration) 혹은 집중(concentration)에 의한 외부효과는 사회복지를 증진시킬 것이다. 특히, 시장에서 양의 외부성을 갖는 경제활동은 사회적으로 최적인 수준보다 적다는 점을 고려하면, 조세는 시장실패 문제를 완화시킨다는 점에서 효율성을 제고할 수 있다.

셋째, 사회적으로 볼 때 바람직하지만 실제로 개인들이 적극적으로 하지 않는 행위, 예를 들어 노후 혹은 주거를 대비한 저축, 혹은 기부 등과 같은 행위를 촉진시키기 위한 방안으로 조세제도를 활용할 수 있다.[6]

따라서 다양한 미시경제적 효율성 측면, 사회적 형평성 증진과 같은 정책목표 달성 측면을 충분히 고려하여 조세정책이 이루어질 필요가 있다.[7] 좀 더 구체적으로 말하면, 소득세제와 관련된 과세정책은 정책대상과 목적, 그리고 그에 따른 경제적 평가 기준에 대한 사전적·사후적 검

---

6) 한편, 정부의 재정지출에 비해 조세지출의 상대적 장점이 존재한다. 조세지출은 정부 개입이 상대적으로 적고, 재정지출에 비해 대상자 선별의 복잡성과 임의성을 배제할 수 있다는 장점을 갖는다. 또한 관리 비용도 상대적으로 적다는 점에서 재정지출에 비해 조세지출이 비용 효율적인 수단(cost-efficient)일 수 있다.

7) 이 외에도 소득세제와 관련하여 중요한 정책적 이슈는 조세제도의 투명성(transparency)과 확실성(certainty)에 대한 것이다. 투명성은 정부의 감독, 납세자의 순응 및 관리비용에 영향을 미친다. 수혜대상 범위가 명확하고 좁을수록 과세로 인한 경제적 왜곡과 그 비용이 축소되는 한편, 비용에 대한 측정이 상대적으로 쉽다. 확실성(certainty)은 특정 조세지출 항목의 지속기간과 실질적인 혜택과 관련된 것으로서, 지속기간이 불분명하거나 서로 다른 조세지출 항목과의 상호작용으로 인한 혜택 정도가 불분명할수록 효율성과 형평성은 저해될 것이다. 좀 더 자세한 내용은 Joint Committee on Taxation(2008)의 pp.62~67 참조.

토가 전제될 필요가 있다.[8)]

이러한 맥락에서 소득세제의 주요 정책 대상은 소비, 저축, 혹은 기부 행위와 같은 납세자의 행태이며, 과세는 이들 관련 행태를 사회적으로 바람직한 방향으로 변화시키기 위한 유인으로 제공될 수 있다. 이와 같은 경우 해당 정책에 대한 평가의 기준은 경제적 효율성 혹은 사회적 가치의 증진이 적절할 수 있다.

한편, 정부는 장애인, 한부모 가구, 혹은 부양자(자녀, 부모, 연령 등)의 수와 구성 등과 같이 담세력(ability to pay)이 상대적으로 부족한 납세자의 세부담을 경감시켜 재분배 혹은 형평성 제고를 직접적인 목적으로 할 수도 있다. 이러한 맥락에서 Avram(2014)은 비록 효율성 손실이 존재한다고 할지라도 소득세제를 통한 형평성 증진은 그 자체로도 정당화될 수 있으며, 특정 정책이 의도했든 의도하지 않았든지 간에 분배상태에 미치는 영향에 대한 검토가 중요하다고 말한다.[9)]

다른 한편, 우리나라의 근로장려세제와 같이 많은 국가에서 시행하고 있는 세액환급(refundable tax credits) 제도는 납세자의 행태(근로유인 등)와 경제・사회적 특성(자녀수 등)이 동시에 정책대상이 된다는 점에서 효율성과 형평성이 동시에 평가될 필요가 있다(TARC, 2014).

하지만 현실에서 소득세제의 사회・경제적 효과를 평가하는 것은 쉽지 않다. 방금 말한 바와 같이, 소득세제가 다양한 측면(개인의 행태, 소득분배 및 세수 등)에 영향을 미친다는 점도 하나의 원인이지만, 무엇보다 소득세제에 대한 평가를 어렵게 하는 요인은 소득세제를 구성하는 요인들이 상당히 복잡하고 상호 연관되어 있기 때문이다.

[그림 4-1]에서 보듯이 대부분 국가들과 마찬가지로 우리나라의 개인소득에 대한 과세체계는 종합과세Global Income Tax)를 원칙으로 하는

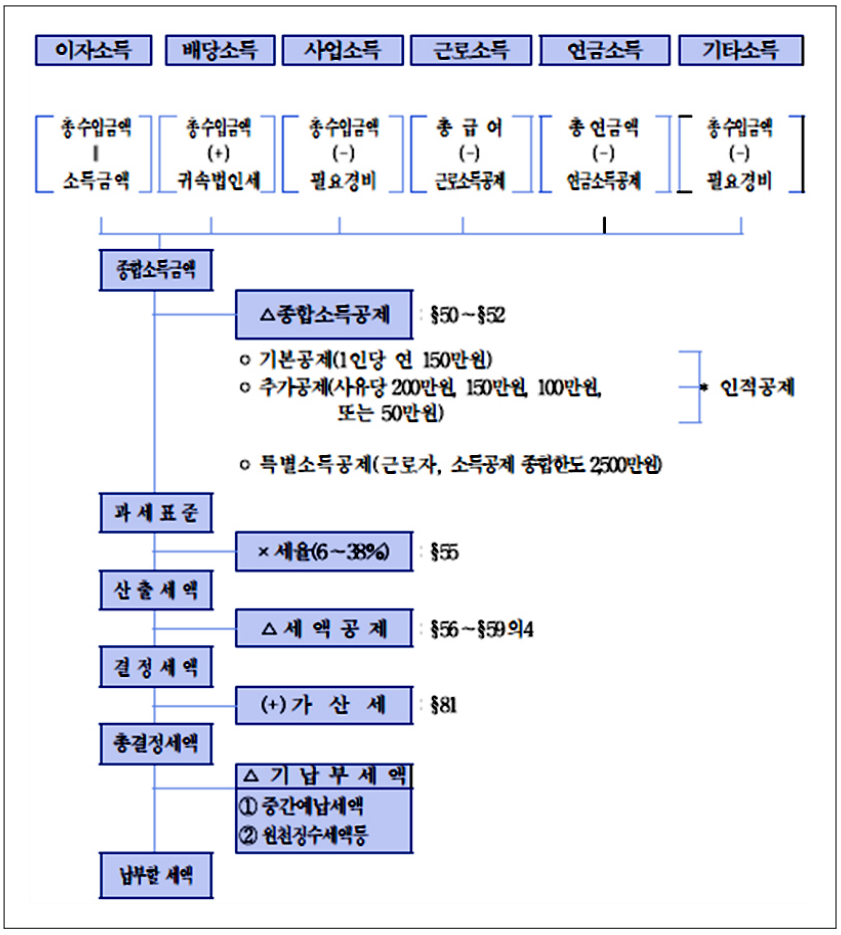
---

8) TARC(2014) 참조.

9) 이러한 맥락에서 소득세제의 재분배효과에 대해 많은 연구가 이루어졌다. 이때, 미국의 소득세제 내 조세지출을 분석하고 있는 연구들 대부분은 조세지출의 역진성을 지적하고 있지만(Howard, 1997; Burman, Geissler et al., 2008; Toder, Harris et al., 2009; CBO, 2013), 최근 유럽 국가들의 조세지출을 비교하고 있는 연구결과들을 보면 국가 혹은 조세지출 유형에 따라 다양한 결과가 나타난다(Avram, 2014; Verbist, 2004; Verbist and Figari, 2014).

데, 이자·배당·기타 소득은 일정금액 이하의 경우 종합소득에 합산하지 않고 분리과세한다(Separate Taxation).[10] 이때, 각 소득원별 필요경비 등을 제외하고 합산한 금액, 즉 종합소득금액에서 다양한 목적과 유형의 종합소득공제를 차감하여 과세표준을 산출한다. 다음으로 과세표준에 누진적 세율(5단계 누진세율체계로 최저세율은 6%, 최고세율은 38% 수준)을 적용하고 세액 공제·감면을 차감하여 최종적으로 세액을 결정한다.

〔그림 4-1〕 우리나라 종합소득과세체계

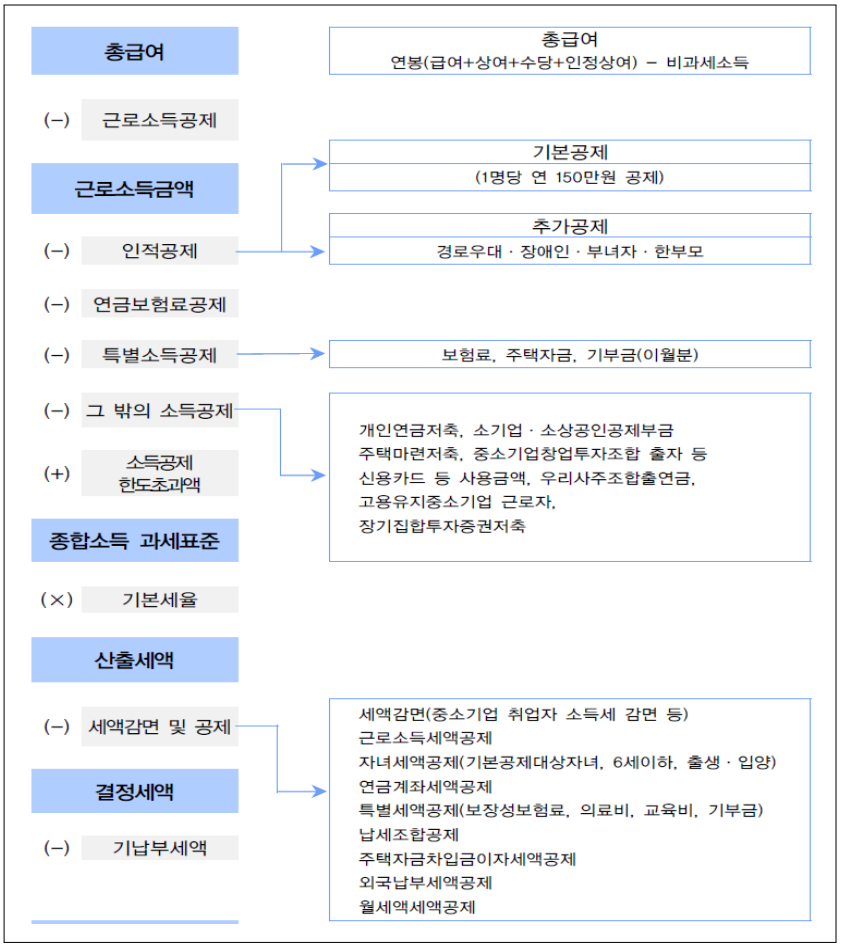


자료 : 기획재정부(2016), p.51.

10) 국회예산정책처(2017)의 일부 내용을 발췌·정리하였음을 밝힌다.

종합소득공제는 인적공제·연금보험료공제·특별공제·조특법상 소득공제로 구성되는데, 인적공제는 납세의무자의 최저생계비에 해당하는 소득을 과세에서 제외시키기 위하여 공제하는 금액이다. 연금보험료공제는 공적연금에 대한 기여금에 대한 공제이며, 특별소득공제는 소득을 창출하기 위한 비용은 아니지만 담세력을 축소하는 지출항목에 대한 공제이며, 조특법상 소득공제는 종합소득의 종류에 따라서 차등 적용된다. 한편, 근로소득만 있는 경우의 세액결정과정을 [그림 4-2]에서 볼 수 있는데,

〔그림 4-2〕 근로소득자의 연말정산 세액계산 흐름도



자료 : 국세청(2016), 「근로자를 위한 연말정산 안내」, p.23.

여기서 근로소득자에게 적용될 수 있는 공제항목들을 좀 더 구체적으로 볼 수 있다.

다시 말하면, 소득세의 누진 세율이 적용되는 과표는 소득금액에서 [그림 4-2]의 다양한 소득공제들을 차감하여 산출된다. 따라서 소득수준이 동일하더라도 개인과 그가 속한 가구의 다양한 인구통계학적 특성과 경제·사회적 행태에 따라 과표가 다를 수 있고, 그 결과 개인이 납부하는 세액 역시 차이가 있다. 즉 다양한 소득공제 항목들은 개인들의 행태 변화와 밀접한 관련이 있는 한계유효세율에 중요한 영향을 미친다. 따라서 소득세제를 통해 개인의 바람직한 행태를 유인하는 정책이 단순히 세율구조의 변화만을 의미하는 것은 아니다. 오히려 좀 더 다양한 정책 목적을 고려한다면 소득공제 제도를 변화시키는 정책이 더욱 효과적일 수 있다.

<표 4-1>에서 보듯이 각종 공제항목들은 나름의 정책적 목적을 갖고 있다는 점에서 소득공제 제도를 활용한 다양한 정책들이 이루어질 수 있다.[11] 예를 들어, 근로자 소득·세액 공제는 '필요 경비성 지출'을 고려한다는 목적 외에도 근로자와 자영자 간의 수평적 형평성 제고라는 목적을 동시에 갖고 있으면서 실제 그로 인한 혜택의 규모도 상당히 큰 편이다. 주거안정을 위해 자가 소유의 비용(주택저당차입금에 대한 이자 등) 혹은 저소득자의 주거비용(월세)을 일부 경감시킬 수도 있다. 또한 다양한 목적의 저축(노후생활, 주택마련 등)을 장려하기 위해 공제제도는 우리나라뿐만 아니라 다른 국가들에서도 흔히 볼 수 있다. 또한 근로빈곤층의 근로유인 제고 및 실질소득 지원, 자녀 양육부담 완화 등을 목적으로 하는 세금 환급형태의 근로연계형 소득지원제도인 근로장려세제(EITC)와 자녀장려세제(CTC)가 2008년과 2015년부터 각각 도입되었고, 이후 그 지급범위와 대상이 점차 확대·조정되고 있다.

한편, 우리나라 소득세제에서 최근 가장 큰 변화는 2014년(귀속소득)부터 이전의 특별공제(의료비, 교육비, 기부금, 보장성보험료), 표준공제, 일부 인적공제(6세 이하, 출산·입양, 다자녀추가공제) 등의 소득공제 제도를 각각 특별세액공제, 표준세액공제, 자녀세액공제 등 세액공제로 전환

11) <표 4-1>은 김태일·박종수(2010)의 개별 소득공제 항목들의 목적에 대한 기본적인 내용을 참조하여 분류한 것이다.

한 것이다. 이것은 납세자의 소득수준이 높을수록 공제혜택이 더 큰 기존의 방식을 개혁한 것으로서 세후소득의 분배상태를 크게 개선시킬 것으로 예상된다. 하지만, 소득수준이 유사하다고 할지라도 납세자 개개인의 다양한 경제·사회적 특성에 따라 그 영향이 상이하게 나타남으로써 납세자들의 불만이 상당히 컸다. 그 결과, 2015년 초 연말정산 당시 해당 세제개혁으로 인한 사회적 논란이 발생했고, 과세당국은 유례없는 소급입법을 시행하기도 하였다. 이것은 소득세제의 경제·사회적 파급효과가 얼마나 큰 것인지를 가늠해 볼 수 있었던 사례인 동시에, 소득세제의 변화가 개인의 세후소득에 미치는 영향에 대한 사전적 분석과 평가가 얼마나 중요한 것인지를 상기시켜 준다.

하지만 소득세제의 경제·사회적 영향은 다양한 요인들의 복잡한 상호과정을 거치게 됨으로써 그 분석과 평가가 쉽지 않은 것이 사실이다. 예를 들어, 실업 혹은 비경제활동 상태의 개인에게 인센티브를 줌으로써 경제활동을 유인한다거나, 또는 경제활동을 하지만 빈곤 혹은 사회적 배제의

〈표 4-1〉 정책목적별 소득공제의 분류

공제 목적	비고
필요 경비성 지출 및 근로자와 자영자 간 수평적 형평성 제고	근로자 소득·세액 공제
최저생활보장	기본 인적공제, 추가 인적공제 (경로우대, 장애인, 부녀자, 한부모)
출산장려	자녀세액공제 (기본공제대상자녀, 6세 이하, 출생·입양)
사회보험 가입 장려 및 비용 경감	국민연금·건강·고용 보험료 근로자 부담분
저축 장려	연금계좌, 개인연금저축, 투자조합출자 등, 장기주식형 저축 등
주거 안정	주택마련저축, 장기주택저당차입금, 주택임차차입금 및 월세액
불가피한 지출 보전	의료비, 교육비
사회적으로 바람직한 지출에 대한 장려	기부금·보장성보험료
근로유인 제고 및 형평성 증진	근로장려금·자녀장려금
상거래 투명성 확보	신용카드 등 사용액

위험에 있는 저소득자를 지원하는 과세정책에 대한 평가에서 정책 시행 전후 한계유효세율(marginal effective tax rates)의 변화에 대한 측정은 필수적이다. 왜냐하면 한계유효세율의 변화는 세제변화에 따른 잠재적 행태변화를 의미하는 것으로서, 정책 대상 집단의 노동공급 변화와 그에 따른 소득의 증감을 추정할 수 있는 주요한 지표이기 때문이다.

하지만 한계유효세율의 변화를 개인 단위로 추정하는 것은 쉽지가 않다. 왜냐하면 세제상의 혜택은 일반적으로 특정 요건에 부합하는 개인을 대상으로 제공되기 때문이다. 예를 들어, 세제혜택은 수급 자격(eligibility)과 관대성(generosity)에 따라 수혜범위와 효과가 달라진다. 근로유인 제고를 위한 세제혜택의 수급자격은 소득수준(개인, 가구, 혹은 개인과 가구 모두), 고용 상태(임금/비임금), 혹은 근로시간에 의해 결정된다. 관대성은 납세자의 상황(소득, 연령, 가구구성, 부양자수 등)에 따른 혜택수준, 수급 가능성, 공제혜택을 받기 위한 대기 시간에 주로 의존한다. 한편 세율감면과 면세(exemptions)에 대한 수급자격은 대부분 소득수준만을 고려한다. 따라서 이러한 제도의 세부적인 특성을 반영하여 개인들이 실제로 받게 되는 세제혜택과 그에 따른 한계유효세율의 변화를 추정하기 위해선 개인·가구의 인구통계학적 특성과 경제·사회적 특성에 대한 정보를 제공하는 미시자료와 그에 따른 방법론 개발이 필수적이다. 만일 『국세통계연보』와 같은 집계자료를 활용하여 소득세제의 정책적 효과를 분석하고자 한다면, 임의적인 가정들과 그에 따른 오차를 피하기 어렵기 때문이다.

그럼에도 불구하고 정부는 조세지출예산서상의 지출규모를 국세통계연보의 과표구간별 납세자수와 납세액 규모, 그리고 각 집단별 실효세율을 활용하여 '과세표준 변화분×실효세율' 혹은 '과세표준×실효세율 변화분'으로 단순하게 추계하고 있다. 좀 더 구체적으로 자녀교육비 공제를 확대하는 법안의 비용추계 방법을 살펴보면, 총 과세표준의 변화분을 기존의 공제금액과 혜택을 받는 영유아~고등학생교육비 비율을 곱하여 산출하고, 산출한 결과에 근로자들의 기존 실효세율을 곱하여 해당 조세지출 규모를 추정한다. 이때, 명목세율이 아닌 유효세율을 곱하는 것은 각종 공제항목들의 영향으로 과세표준과 명목세율의 곱이 그대로 세액이

되지 않는다는 점을 고려한 것으로 보이지만, 이를 위해선 세제변화가 세액계산에 필요한 항목들(소득, 소득공제, 세액공제 등)에 미치는 영향이 개인단위로 측정되어야 한다.[12)]

한편, 김태일・박종수(2010)는 국세통계연보의 소득구간별 소득액, 과세액, 세부 항목별 공제액 집계치를 활용하여 소득세 조세지출의 형평성을 분석하고 있다. 이를 위해 국세통계연보의 자료를 나름의 가정하에 10분위 소득계층으로 전환하고, 각종 공제액의 집중계수를 산출한다. 이것은 10명의 개인이 전체 임금근로자들을 대표한다고 가정하는 것으로서 그에 따른 편의(bias)를 피할 수 없다. 예를 들어, 동일한 소득계층에 속한다고 할지라도 개인마다 받고 있는 공제혜택은 다를 수 있지만, 집계자료를 활용하는 경우에는 이를 고려할 수 없다. 또한 2015년 연말정산 파동의 원인과 대안을 제시하고 있는 최보람(2015)에서도 국세통계연보상의 총급여 구간별 총 금액을 각 구간의 과세대상인원으로 나누어 1인당 금액으로 소득계층을 구분하고, 근로소득금액, 과세표준, 산출세액 및 결정세액 역시 1인당 금액으로 환산하여, 소득공제의 세액공제 전환에 따른 각 계층별 한계세율의 변화를 추정하고 있다. 따라서 동 연구 역시 동일한 소득계층에 속한 개인들이 받는 공제혜택들이 유사 혹은 동일하다는 비현실적인 가정이 전제될 수밖에 없다.

정리하면, 소득공제와 같은 조세지출은 일반적으로 특정 집단을 대상으로 한다는 점, 실효세율이 소득수준에 따라서 다를 수 있다는 점, 그리고 소득공제의 혜택은 개인의 경제・사회적 특성에 의존한다는 점을 고려하면 집계자료를 활용한 추정방법의 한계는 명확하다. 다시 말하면, 소득세는 소득금액에서 소득공제를 제외한 과표를 기준으로 누진적인 명목세율로 산출세액을 산출하고, 산출세액에서 세액공제를 차감하여 최종적으로 세액을 결정한다. 따라서 소득공제항목을 포함한 세제개편으로 인해 과표가 변화한다면, 이전과는 다른 명목세율이 적용되는 납세자의 세부담 변화를 포착할 필요가 있다. 즉 소득공제 → 과표 ⇒ 적용 세율 ⇒ 산출세액 ⇒ 세액공제[13)]로 이어지는 경로를 고려한 분석이 필수적이다. 하

12) 국회예산정책처(2012)는 행정부가 조세지출에 대한 비용 추계에 사용한 가정과 자료의 출처, 정의 등이 제시되지 않는 문제점을 지적하고 있다.

지만 집계자료를 이용한 분석방법은 이러한 경로를 고려할 수 없기 때문에 세제변화에 따른 세수와 조세지출의 규모 변화를 과대 혹은 과소평가하는 편의를 피할 수 없다. 더욱더 큰 문제는 세제 변화가 모든 개인에게 동일한 방향이 아닌, 서로 다른 방향으로 영향을 줄 수 있음에도 불구하고, 집계자료를 이용한 분석은 이를 포착·분석하지 못함으로써 세제의 형평성 분석에도 적지 않은 문제를 발생시킬 수 있다.

이러한 맥락에서 최근 많은 국가가 소득세제가 미치는 경제·사회적 영향을 분석하기 위해 개인·가구의 경제·사회적 속성을 조사·제공하는 미시자료(횡단면 혹은 패널)에 기반하여 소득세추정모형을 개발·활용하고 있다는 점에 주목할 필요가 있다.[14] 예를 들어, Barrios et al. (2015)는 5개 유럽국가(프랑스, 스페인, 영국, 헝가리, 슬로바키아)에 대한 연구에서 EUROMOD를 활용하여 필요한 모수들을 추정한다. CBO (2012)에서도 중·하위 소득 계층의 한계유효세율 변화를 CBOLT를 활용·분석하고 있다. 한편, EC(2014)에서는 많은 국가에서 공통적으로 시행되고 있는 근로유인 제고, 주거지원 및 저축 등을 목적으로 하는 세제정책들의 경제·사회적 효과를 분석하기 위한 실증분석 방법론에 대한 논의들을 정리·제시하고 있다는 점에서 관련 연구에 유용한 참고가 될 것이다.

일반적으로 미시자료는 표본으로 선정된 일정수의 가구나 개인들의 거

---

13) 예를 들어, 2015년 기준 근로소득세액공제는 근로소득에 대해 적용되는 세액공제로 산출세액이 130만 원 이하일 경우 55%, 130만 원을 초과하면 30%를 공제한다. 공제한도는 총급여액이 3,300만 원 이하면 74만 원, 7,000만 원 이하는 74만 원 - (총급여 - 3,300만 원) * 0.008, 7,000만 원 초과는 66만 원 - (총급여 - 7,000만 원) * 1/2이다.

14) 행정부와 국회예산정책처 역시 조세정책 효과 분석을 위한 미시모의실험모형의 개발·활용의 필요성을 인지하고 있다. 예를 들어, 국회예산정책처(2012)와 박명호 외 2인(2008)에서는 미시모의실험모형을 통한 조세지출 비용 추정을 시도함으로써 향후 그 활용가능성을 모색한 바 있다. 예정처(2012)는 통계청의 「가계동향조사」자료를 이용하여 이동통신요금 지출액에 대한 소득공제 제도 신설의 효과에 미시모의실험모형을 활용하였다. 이때 행태변화가 없다는 가정하에 세수변화를 추정함으로써 세수손실법에 의한 비용 추정만을 시도하였다. 박명호 외 2인(2008)은 「재정패널」자료를 이용하여 보험료 소득공제의 세수비용 추정 사례를 통해서 미시모의실험모형, 분포모형 및 기존 방법의 장·단점을 비교·검토한다.

주특성, 소득, 지출상황 등 조세지출 비용 추계에 필요한 정보를 포함하고 있다. 따라서 특정 조세지출의 대상자 규모와 특성별 분포를 파악하는 것이 가능하다. 이에 미시자료를 활용한 시뮬레이션 분석은 표본 내 개인들의 경제·사회적 특성(소득, 공제관련 지출 내역 및 항목 등)을 바탕으로 정책 변경 전후의 세부담 변화를 납세자 개인 단위로 산출할 수 있다. 나아가 미시자료가 일반적으로 제공하는 모집단 가중치를 적용함으로써 세제개혁에 따른 세수 추계 및 재분배 효과를 분석할 수 있다. 나아가 전체 소득세 체계를 구성하는 세부 요인들의 변화에 따른 효과에 대한 분석이 가능함으로써, 다양한 목적에 따른 세제정책의 효과를 분석할 수 있다. 특히, 국세청의 개인소득세 신고자료 표본을 기반으로 한다면 추정결과의 정확도와 신뢰성이 크게 제고될 것이다.[15] 하지만 아직까지 우리나라에서는 개별 납세자료가 추계 목적으로 제공되지 않고 있는 실정이어서, 가계동향조사(통계청), 재정패널(한국조세재정연구원), 노동패널(한국노동연구원) 혹은 복지패널(한국보건사회연구원) 등과 같은 미시 패널 자료를 활용한 노력이 이루어지고 있다.[16]

하지만 미시자료가 있다고 해서 모든 문제가 해결되는 것은 아니며, 이를 활용한 분석방법론에 대한 연구와 소득세 관련 모형 개발이 필요하다. 왜냐하면 다른 패널자료에 비해 개인의 납세 관련 정보가 상대적으로 많은 재정패널 자료조차도 각종 세법 변화에 따른 효과를 분석하는 데 필요한 정보가 부족하기 때문이다. 또한 대부분의 국내 미시자료들에서는 연간 근로소득이 1억 원 이상인 고소득자에 대한 표본이 상대적으로 매우 적은 문제가 있다. 예를 들어, 2015년 근로소득세수 약 25.4조 원 중에서 근로소득금액이 1억 원 이상 계층이 납부하는 세액 비중은 약 42%(약

15) 소득세 신고자료를 표본으로 활용하고 있는 국가로는 캐나다, 호주 및 미국이 대표적이다. 자세한 내용은 박명호·김우철·전병힐(2008)과 국회예산정책처(2012) 참조.

16) 국회예산정책처(2012)와 박명호 외 2인(2008)에서는 미시모의실험모형을 통한 조세지출 비용 추정을 시도함으로써 향후 그 활용가능성을 모색한 바 있다. 예정처(2012)는 통계청의 「가계동향조사」자료를 이용하여 이동통신요금 지출액에 대한 소득공제 제도 신설의 효과에 미시모의실험모형을 활용하였다. 박명호 외 2인(2008)은 「재정패널」자료를 이용하여 보험료 소득공제의 세수비용 추정 사례를 통해서 미시모의실험모형, 분포모형 및 기존 방법의 장·단점을 비교·검토한다.

10.8조 원)에 이른다. 마지막으로, 미시자료는 조사대상 개인의 응답에 의존한다는 점에서 추정결과의 해석에 유의할 필요가 있다. 나아가 실적치와의 비교를 통해 추정결과의 편의를 수정・개선할 수 있는 분석방법에 대한 연구가 필요하다. 이러한 맥락에서 미시 모의실험 모형을 개발하거나 활용하고 있는 기존 국내・외 연구들을 살펴보는 것은 향후 연구발전에 유용할 것이다.

2. 국내・외 사례 및 시사점

가. 해외 사례 : EUROMOD

Orcutt(1957)은 거시 모형의 한계를 지적하면서, 개인, 가족, 그리고 기업과 같은 미시적 의사결정주체 간 상호작용을 중심으로 하는 새로운 분석모형으로 미시 모의실험 모형을 이용한 분석방법을 처음으로 제안하였다. 하지만 미시 모의실험 모형이 본격적으로 발전하게 된 시기는 미시자료의 축적, 컴퓨터의 발전, 그리고 사회・경제정책(조세・연금・교육・복지 등)의 장기적인 효과에 대한 정책적 수요에 따른 재정적 지원이 이루어지게 된 1990년대 초반부터이다. 이후 정부기관과 연구소를 중심으로 개발되었거나 활용되고 있는 동학적 미시 모의실험 모형들은 18개 국가에서 60개에 이르며, 모형의 주요 분석 대상은 연금, 건강, 조세 및 기타 복지 등으로 다양하다(Li and O'Donoghue, 2013).[17]

미시 모의실험 모형은 시간개념과 행태변화 포함 여부에 따라 크게 정태(static) 모형과 동학(dynamic) 모형으로 구분되는데, 조세/복지 모형(tax/benefit model)은 대부분 정태적 모형이다. 정태 모형은 정책 이외 요소들의 변화가 없다는 것을 전제로 조세/복지 정책의 효과를 분석한다.[18] 예를 들어, EU 28개 국가의 조세/복지 정책에 대한 모형으로 다른

17) 일반적으로 미시 모의실험 모형은 모형 내에 시간개념과 행태변화 포함 여부에 따라 크게 정태적(static) 모형과 동학적(dynamic) 모형으로 구분한다. 이때, Li and O'Donoghue(2013)의 주요 내용은 동학적 모형들의 방법론적 이슈에 따라 모형들을 분류・정리하고 있다. 한편, 국내 개발 사례로는 고제이・권혁진 외 4인(2016)이 있다.

모형보다 상대적으로 널리 알려진 EUROMOD는 정태 모형이며, 영국의 TAXBEN(Brewer et al., 2007), 스웨덴의 SWEtaxben(Ericson et al., 2009), 독일의 IZAΨMOD(Peichl et al., 2010), 룩셈부르크의 LuxTax Ben(Islam and Flood, 2015) 및 OECD의 Tax-Benefit model(TBM) 등도 정태 모형이다. 동학 모형으로는 미국 CBO의 CBOLT, 영국의 SAGE (Zaidi and Rake, 2001; Zaidi and Scott, 2001), 스웨덴의 MICROHUS (Klevmarken, 1991; Klevmarken and Olovsson, 1996) 및 독일의 MICSIM (Merz et al., 2002) 등이 있다. 국내에서는 정태적 모형으로 KIPFSIM(성명재 외 2인, 2008, 2010)의 개발이 이루어진 적이 있지만, 이후 해당 모형을 활용한 연구를 찾아보기 힘든 것은 아쉬운 점이다. 한편, 소득세 모듈이 포함된 동학 모형으로는 DOSA(고제이・권혁진 외 4인, 2016)가 있지만, 해당 모형은 연금제도에 대한 평가가 주요 목적이기에 소득세제에 대한 평가 목적으로 활용되기 위해선 향후 모형의 개선과 발전이 필요하다.

한편, 이상의 많은 모형 중에서도 소득세추정모형에 대한 개발과 활용에 있어서 주목할 만한 모형은 EUROMOD이다. 해당 모형은 모형의 기술적 특징과 적용 분야 등에 대한 정보가 정책담당자들과 연구자들은 물론, 일반인들에게도 광범위하게 공개되는 등 현재 가장 폭넓게 이용되고 있는 모형이기 때문이다. 이에 EUROMOD에 대해 좀 더 자세히 살펴보는 것은 매우 유용할 것으로 판단된다.[19]

EUROMOD는 유럽연합(EU)의 조세/복지 모형(tax/benefit model)으로서 각 회원국가의 조세/복지 제도와 정책에 따라 개인・가구의 납세액 및 복지혜택을 시뮬레이션 하는 모형이다. 기술적인 측면에서 보면, EURO MOD는 개인의 행태와 인구의 사회-인구통계학적 특성이 고정되어 있다는 가정하에서 개인・가구의 조세/복지혜택을 각 국가별 미시자료에 기반하여 계산(arithmetic simulation)한다는 점에서 정태적 미시 시뮬레

18) 이에 해당 모형을 'one night effect' 혹은 'morning after effect' 모형으로 부르기도 한다.

19) EUROMOD에 대한 내용은 Sutherland and Figari(2013)과 Sutherland(2016)의 내용을 요약・발췌하는 한편, 홈페이지(www.euromod.ac.uk)의 내용을 참조하여 서술한 것임을 밝힌다. 이에 별도의 인용이 필요한 경우가 아니라면, 관련 인용들은 생략하였음을 양해 바란다.

이션 모형(static microsimulation model)이다.

Sutherland(2013)의 연구에 따르면, EUROMOD의 기본적인 아이디어는 Atkinson et al.(1988)의 연구에서 비롯됐다. 동 연구는 당시 프랑스의 개인·가구의 인구통계학적 특성과 경제·사회적 특성에 대한 정보를 제공하는 미시자료, 즉 프랑스의 인구 구성과 경제·사회적 상황에 프랑스의 제도가 아닌 영국의 조세제도와 가족수당을 적용하는 경우의 분배상태와 빈곤의 변화를 분석하였다. 이후 Sutherland(1997)에서 EUROMOD의 원형(prototype)이 소개되었으며, 이를 활용한 Callan and Sutherland (1997)에서는 영국과 아일랜드를 대상으로 기본소득이 분배상태에 미치는 영향을 분석하였다. 사실, 동 연구에서는 다른 많은 유럽국가도 분석대상에 포함시키고자 하였지만 개별 국가에 대한 정보가 당시에는 부족하여 영국과 아일랜드만을 분석한 것이다. 하지만, 동 연구는 각 국가의 고유한 경제·사회적 상황과 법·제도적 환경을 고려하면서도, 조세/복지 정책의 국가 간 비교가능성을 높이기 위한 일관된 방법론 개발에 대한 필요성을 인식하게 되는 계기가 되었다.

이에 EC의 지속적인 재정적 지원하에 각 국가의 대학, 연구기관 및 정부 부처 소속 연구자들이 참여하여 공동 연구를 수행하는 한편, 관련 전문지식과 자료를 구축하게 되었다. 모형 개발과 유지, 참여자들의 업무조정 등은 초기에는 Sutherland 교수를 중심으로 캠브리지 대학의 연구자들이 맡았지만, 현재는 Essex대학의 사회경제연구소(ISER)에서 맡고 있다.

나아가 ISER과 28개 국가의 관련 담당자들 간 네트워크를 통해 매년 업데이트를 수행하고 있고 그 결과를 공유하고 있다. 또한 정기적인 교육과정을 통해 모형을 활용하는 연구자들과 정책담당자들이 늘어나고 있고, EU 소속이 아닌 국가들의 모형 개발을 지원하기도 한다.

현재 EUROMOD(공개버전 G4.0)[20]는 EU 소속 28개 모든 국가의 조세

20) 초기에는 Excel 매크로 형태로 제공되었지만, 현재는 C++로 컴파일된 소프트웨어를 사용하여 인터페이스와 분석을 위한 많은 옵션이 사용자에게 편리하게 이루어져 있다. 예를 들어, 모수적인 차원에서 논의되는 제도 개혁의 효과 정도는 사용자가 별도로 프로그램을 작성하지 않고도 산출할 수 있다. 한편, 해당 모델을 가능한 한 최신 상태로 유지하기 위한 각 국가의 정책변화 내용과 입력데이터의

/복지제도에 대한 2013~2016년 시뮬레이션 결과를 산출·제공하고 있다.[21] 좀 더 구체적으로 말하면, 다른 미시 모형과 마찬가지로 EUROMOD는 각 국가별 조세/복지 정책과 제도에 따른 총소득·순소득 등과 같은 기본적인 결과를 계산하는 한편, 해당 정책과 제도의 변화가 소득분배, 정부예산 및 개인의 경제·사회적 유인(예를 들어, 근로와 저축 등)에 미친 영향을 분석한다. 또한 여러 시점에 대한 분석을 통해 공공 정책의 변화가 소득 빈곤이나 불평등을 감소(또는 증가)시키는 데 기여한 정도를 비교·평가한다. 나아가, 국가별 및 EU 수준에서 제안·논의되고 있는 경제 또는 인구 통계학적 시나리오 혹은 정책적 대안에 따른 영향을 시뮬레이션을 통해 비교·검토한다.

EUROMOD를 활용한 학술적 연구와 정책적 활용사례는 지난 10년 동안 크게 증가·발전했다.[22] 특히, 최근에는 EUROMOD가 국제기구의 주요 정책 분석 도구 중 하나로 활용되고 있다는 점에 주목할 만하다. 예를 들어, EC의 Directorate General for Employment, Social Affairs and Inclusion(DG-EMPL)은 '고용 및 사회 발전에 대한 연례 보고서(Annual Report on Employment and Social Development)'와 'EU Employment and Soical Situation Quartrly Review'에 EUROMOD의 분석결과를 정기적으로 포함시키고 있을 뿐만 아니라, 회원국들 간 토론에도 기초자료로 활용하고 있다. 또한 EC의 다른 정책분과(예를 들어, Taxation and Customs Union(DG-TAXUD), Economic and Financial Affairs(DG-ECFIN), Justice(DG-JUST))에서도 성평등과 주택에 대한 과세 등과 같은 다양한 분야에 대한 EUROMOD의 분석결과를 활용하고 있다.

좀 더 구체적으로 말하면, DG-EMPL에서는 현재의 빈곤위험을 평가

업데이트 등은 유럽집행위원회(European Commission)의 자금 지원으로 이루어지고 있다.

21) 2005년 모형은 9개 국가, 2006년은 18개 국가, 2007년 27개 국가, 2011년부터 28개 국가가 포함되어 있다. 한편, Sutherland(2013)는 EUROMOD가 오스트리아, 그리스, 라트비아, 룩셈부르크, 몰타, 포르투갈, 스페인 등과 같이 조세/복지 모형을 통한 분석이 일반적으로 이용 가능하지 않았거나 존재하지 않았던 국가에서 관련 모형을 개발하려는 자극제가 되었다고 평가하고 있다.

22) EUROMOD를 활용한 연구, 프로젝트, 정책 보고서 및 활용사례 등은 홈페이지(https://www.euromod.ac.uk/research) 참고.

하기 위해 EUROMOD의 시뮬레이션 결과를 활용한다. 왜냐하면 특정연도에 조사・수집된 서베이 조사는 일반적으로 1~2년 전의 소득을 조사하는 경우가 많기 때문이다. 또한 매년 논의되고 있는 재정 정책의 개혁이 분배상태에 미치는 효과를 사전적으로 검토하기 위해 EUROMOD를 활용하고 있다. 특히, 최근 경제위기와 이후 극복과정에서 각 국가의 재정 정책이 분배상태에 미치는 영향에 대한 연구들과 EU 혹은 EMU 차원에서 통합적인 사회・재정정책의 설계와 안정화를 평가하기 위한 연구들에서 EUROMOD가 주요 분석도구로 활용되고 있다.

다른 한편, EC 이외의 국제기구들에서도 EUROMOD에 대한 관심과 실제 활용사례들이 점차 증가하고 있다. IMF(2014)의 경우 재정정책이 소득불평등에 미치는 영향에 대한 주요 연구로 EUROMOD를 활용한 5편의 연구를 참고하는 한편, 긴축정책이 분배상태에 미치는 영향에 대해 EUROMOD를 활용하여 직접적으로 추정・분석한 바 있다.[23] 세계은행(World Bank)은 발칸국가들에 대한 EUROMOD의 확장에 관심을 표명하기도 하였으며, UNICEF는 어린이를 대상으로 한 복지혜택에 대한 EUROMOD의 연구결과에 주목한 바 있다.[24] 그리고 그리스에 관한 최근 OECD 보고서도 EUROMOD를 적극적으로 활용하고 있다.[25]

다음으로 EUROMOD의 좀 더 세부적인 기술적인 내용은 다음과 같다. EUROMOD에 활용되는 데이터는 기본적으로 Eurostat에 의해 발표되는 '소득 및 생활 조건에 관한 유럽연합 통계(EU Statistics on Income and Living Conditions : EU-SILC)' 자료이며, 해당 자료에서 필요한 변수들을 추출・가공한다. 한편, 해당 국가의 통계기관에서 제공하는 SILC에 좀 더 세부적인 정보가 있는 경우에는 해당 자료를 직접 사용하거나 EU 버전의 자료를 보완하기 위해 사용한다(벨기에, 불가리아, 체코, 그리스, 스

23) www.imf.org/external/np/pp/eng/2014/012314.pdf

24) Levy, H., M. Matsaganis, and H. Sutherland(2014), "Simulating the costs and benefits of a Europewide basic income scheme for children", UNICEF Child Poverty Insights, www.unicef.org/socialpolicy/index_71776.html

25) Koutsogeorgopoulou, V., M. Matsaganis, C. Leventi, and J. D. Schneider(2014), "Fairly Sharing the Social Impact of the Crisis in Greece", OECD Economics Department Working Papers No 1106, www.oecd-ilibrary.org/economics/fairly-sharing-the-social-impact-of-the-crisis-in-greece_5jzb6vwk338x-en

페인, 이탈리아, 리투아니아, 룩셈부르크, 오스트리아, 폴란드 및 슬로바키아). 이처럼 EU-SILC이라는 공통적인 자료의 활용은 국가 간 비교 연구 시 데이터의 이질성에 따른 연구의 한계를 극복하는 동시에, 많은 국가에 대한 미시 시뮬레이션 모형 개발과 유지 차원에서 보면 상당히 효율적이라고 평가할 수 있다.

한편, EU-SILC의 단점은 다른 일반적인 서베이 자료들과 유사하다. 예를 들어, 자본소득에 대한 과세의 영향을 분석하는 데 있어서 금융소득에 대한 정보는 상당히 중요함에도 불구하고 그 조사범위는 상대적으로 좁은 편이다. 또한 자산에 대한 정보가 부족한 점과 연간단위로 소득을 환산하는 것이 때때로 어려운 경우가 있는데, 이는 자산-소득조사 방식의 사회부조 프로그램을 시뮬레이션 하는 경우 문제가 될 수 있다. 또한 SILC의 Eurostat 버전은 일부 소득원에 대해서 제도・정책별로 제공되는 것이 아니라, 예를 들어 실업 혹은 노령 등과 같이 기능별로 합산된 금액이 조사・제공된다. 따라서 EUROMOD의 기초자료(input database)를 구성하는 경우 일부 소득원에 대한 의제삽입(imputation)과 그에 따른 추정의 정확성 감소는 불가피하다(Figari et al., 2007).

EUROMOD의 기초자료는 개인단위로 제공되는데, 개인이 속한 가구의 인구통계학적 특성, 개인의 노동시장 특성, 총 시장소득 및 기타 소득원(즉 연금, 공・사적 이전소득)에 대한 변수들로 구성된다. 모집단에 대한 추정을 위해 개인 가중치 변수가 포함되어 있고, 비응답 및 표본 추출 오류를 수정하기 위해 해당 가중치가 사용되기도 한다. 한편, 설문 조사 중에 총소득(gross income) 금액이 조사되지 않고 부적절한 값이 입력된 경우, 순소득(net income) 변수를 활용하여 총소득을 추정하기도 한다(net-to-gross procedure).

그리고 EUROMOD는 기초자료와 산출결과에서 활용하는 변수들에 대한 정의와 변수명을 표준화함으로써, 국가별로 유사한 조세/복지혜택임을 연구자가 쉽게 파악할 수 있도록 사용자 편의성을 높이고 있다. 나아가 이러한 표준화는 다른 자료들을 EUROMOD의 기초자료로 활용할 수 있는 가능성을 넓힌다는 점에서 모형의 확장성에도 유용하다.

EUROMOD는 개인・가구의 세금 및 복지수당을 시뮬레이션 하여 가

처분소득을 산출하는 것을 목표로 한다. 좀 더 구체적으로 말하면, 소득세(중앙 및 지방), 사회보험에 대한 기여금(근로자, 자영업 및 고용주가 지급), 가족 수당, 주택 수당, 사회 보조 및 기타 소득 관련 수당을 시뮬레이션 한다. 한편, 모형을 통해 시뮬레이션 할 수 없는 것은 서베이 자료에서 제공하는 내용을 가처분소득 산출과정에 적용한다. 예를 들어, 과거의 고용 및 기여 내역에 대한 정보가 필요한 연금, 그리고 장애의 특성과 심각성에 대한 정보를 전제로 하는 장애급여는 조사된 금액을 그대로 활용한다.

나아가, EUROMOD는 '기준선(baseline)' 시뮬레이션을 통해 모형의 유효성을 검증하고 개선함으로써 동 모형에 대한 이용가능성과 연구결과의 신뢰성을 제고하고 있다. 이때 기준선(baseline) 시뮬레이션은 조사된 소득의 귀속연도를 기준으로 이루어진다. 예를 들어, 2007년에 대한 기준선 시뮬레이션은 2007년 소득을 조사・제공하는 2008 EU-SILC 데이터를 활용하여 이루어진다. 그리고 기준선(baseline) 시뮬레이션에 대한 유효성(validation)을 미시와 거시 차원에서 각각 검토한다. 미시적 차원에서는 몇몇 가구를 임의적으로 선택하여 이들 가구에 대한 복지급여 수급 여부와 금액, 그리고 납세액에 대한 시뮬레이션이 적절하게 이루어졌는가를 검토한다. 그리고 개인별로 자료에서 조사・입력된 값과 시뮬레이션 값을 비교하고, 기초통계와 이상치를 확인함으로써 시뮬레이션이 적절하게 이루어졌는가를 확인한다. 거시적 측면에서는 복지급여별 수급자수와 총지출액, 그리고 납세자수와 총세수를 정부부처의 실적 통계 등과 같은 외부 통계와 비교한다. 또한 빈곤율과 불평등도와 같은 소득분배와 관련된 시뮬레이션 결과는 조사・입력된 값을 기준으로 산출한 값과 비교하거나 다른 연구의 결과와 비교한다.

이러한 유효성 검토 결과는 국가별 보고서(Country Reports)를 통해 홈페이지에 공개된다. 한편, 동 보고서에는 유효성 검토 결과 외에도 조세/복지 제도에 대한 기본적인 설명과 더불어 시뮬레이션 대상 제도에 대한 상세한 설명을 포함하고 있다. 또한 국가별 기초자료에 대한 정보(표본 특성, 가중치, 데이터 조정, 의제삽입한 내용과 시뮬레이션 가정 등)도 포함되어 있다. 이것은 EUROMOD를 활용한 연구 결과들의 신뢰성을 높이는 동시에, 많은 연구자의 접근과 이용가능성을 제고한다는 점에서 시

사하는 바가 크다.

다음으로 EUROMOD를 활용한 다양한 학술·정책적 연구들 중에서 주요 분야의 연구들과 그 성과들을 소개하면 다음과 같다.

첫째, 조세/복지 제도·정책의 재분배효과에 대한 연구들이 있다. 특히, EU-SILC에 순소득과 총소득이 조사·제공되기 시작한 2009년 이전 시기에 조세/복지 정책의 재분배효과에 대한 연구에서 EUROMOD의 기여는 상당히 중요했다. 왜냐하면 유럽국가들을 대상으로 정부 정책 이전과 이후의 소득을 제공하는 미시 자료는 EUROMOD 이외에는 거의 존재하지 않았기 때문이다. 예를 들어, Immervoll et al.(2006)은 유럽국가들에서의 세금과 혜택의 재분배효과에 관련하여 가장 많이 인용된 실증 연구 중 하나이며, Paulus et al.(2010)은 현물 혜택(교육, 건강, 공공 주택)을, Decoster et al.(2010)은 간접세를 고려하여 정부 정책의 재분배효과를 검토하였다.

특히, 2007년 금융위기를 극복하는 과정에서 각 국가에서 시행된 재정정책이 소득분배상태에 미치는 영향을 분석하는 많은 연구가 EUROMOD를 활용하였다. 이를 보면 EUROMOD가, 경제위기와 그에 따른 재정정책이 소득분배에 미치는 영향에 대한 최근의 사전적 평가 연구에서 가장 중요한 분석도구임을 어렵지 않게 추론할 수 있다.

둘째, EUROMOD는 유럽 전역 혹은 여러 시기에 걸쳐 나타나는 특정 조세/복지 정책의 재분배효과를 미시적 수준에서 분석할 수 있다는 점에서 매우 유용하다. 예를 들어, Bargain은 EUROMOD를 활용한 일련의 연구를 통해서 조세/복지 정책의 효과를 조세/복지제도의 변화, 인구구조의 변화, 그리고 시장소득분포의 변화에 따른 효과로 분해함으로써 각각의 상대적 효과를 수치화하는 방법론을 개발·확장했다(Bargain and Callan, 2010; Bargain 2012a, 2012b).

셋째, EUROMOD를 활용한 국가 간 정책 스와핑(policy swapping) 분석이 활발하게 이루어지고 있다.[26] 영국의 세제를 프랑스에 적용한 Atkinson et al.(1988)의 연구가 EUROMOD 개발에 하나의 동기가 된 것에서 알 수 있듯이, EUROMOD를 활용한 국가 간 비교 연구를 통해 특정

26) 관련 국내 연구로는 권혁진(2007) 참조.

인구집단에 대한 정책의 영향, 조세/복지제도와 인구의 특성 사이의 상호작용 및 국가 정책의 효율성에 대한 다양한 연구들이 수행되고 있다. 예를 들어, Canto et al.(2012)와 Levy et al.(2007, 2009)에서는 아동복지 정책의 효과를 EUROMOD를 활용한 정책스와핑 분석을 통해 지출수준, 정책과 제도, 그리고 인구구조의 특성에 따른 효과로 비교・평가하고 있다.

넷째, 소득세제의 세부 요소들의 개혁에 따른 전망 및 평가와 관련된 많은 실증 연구가 이루어졌다. 예를 들어, Paulus and Peichl(2009)을 비롯한 많은 연구가 아직까지는 동유럽국가에 한정된 것이기는 하지만, 단일 세율 소득과세가 재분배와 근로유인에 미치는 효과에 대해 실증분석한 바 있다. 또한 최근에는 금융위기 이후 주거에 대한 과세(housing taxation)가 많은 관심을 끌고 있는데, 이러한 맥락에서 Figari et al.(2012b)의 연구는 주목할 만하다. 동 연구는 주거에 대한 과세가 조세수입을 증가하는 동시에 노동소득에 대한 조세격차(tax wedge)를 줄이는 방안임을 분석・제시한다. 또한 Maestri(2012)는 자가 소유의 편익을 고려하여 주거관련 정책들의 재분배효과를 통합하여 분석함으로써 기존 연구들을 확장・발전시켰다는 평가를 받고 있다.

다섯째, 조세/복지제도가 근로유인과 노동공급에 미치는 영향은 학술적・정책적으로 모두 관심이 있는 전통적인 주제라고 할 수 있다. 이때, 대부분의 기존 연구들은 방법론상의 한계로 인해 정책변화가 근로유인 등에는 영향을 주지 않는다는 전제하에서 조세/복지제도 개혁이 재분배 및 재정에 미치는 영향을 분석하였다(일명 overnight effect 분석).

그러나 최근 EUROMOD를 활용한 많은 연구(Immervoll, 2004; Jara and Tumino, 2013; Immervoll and O'Donoghue, 2004; Barrios et al., 2015)는 정책변화가 근로 인센티브를 비롯한 2차 효과(second-round effect)에 미치는 영향에 대한 분석까지 그 범위를 확장・발전시키고 있다. 예를 들어, 근로유인과 관련하여 세액증가 혹은 복지혜택 축소로 인한 노동시간 연장 혹은 노동강도 증가(intensive (hours) margin), 그리고 노동시장 참여 여부(extensive (participation) margin)에 미치는 영향을 EUROMOD를 활용하여 추정한 한계유효세율을 통해 구분・분석하였다.

또한 정책변화에 따른 개인의 행태 변화를 예측하기 위해 잘 알려진

구조적 이산선택 노동공급 모형(structural discrete choice labor supply model)을 활용한 분석에서 필요한 예산집합을 산출하기 위해 EUROMOD를 이용한 연구들이 늘어나고 있다(Bargain et al., 2011, 2012a; Immervoll et al., 2007, 2011). 특히, Bargain et al.(2012c)는 7개 유럽국가의 성별・혼인상태별 노동공급 탄력성을 EUROMOD를 통해 비교・분석하였는데, 동 연구는 서로 다른 방법론과 자료를 활용한 기존 연구들에서 나타나는 측정오차의 문제를 극복한 것으로 평가된다. 정리하면, 일관된 방법론과 자료를 활용하여 추정된 노동공급에 대한 선호도와 탄력성은 조세 제도와 복지 체제의 최적화에 대한 국가 간 비교 연구에서 중요한 지표라고 할 수 있는데, EUROMOD를 통해 이러한 지표를 산출・활용할 수 있다는 점에서 EUROMOD의 기여도를 높이 평가할 수 있다.

나. 국내 사례 : 노동패널과 재정패널의 활용

소득세 제도에 대한 미시 모의실험 모형의 개발과 활용과 관련하여 아직까지 국내에서는 특정 기관이 모형을 개발하고 이후에도 지속적으로 관리・활용한 사례는 없다고 평가할 수 있다. 이에 국내 사례는 노동패널 혹은 재정패널 자료를 활용하여 소득세제의 경제・사회적 효과를 평가한 연구들을 살펴보고자 한다. 좀 더 구체적으로 말하면, 이하에서는 노동패널을 활용한 6편의 연구들[27]과 재정패널을 활용한 18편의 연구들[28]에 대해 그 방법론적 측면을 중심으로 살펴보고자 한다.

구체적인 사례검토에 앞서, 노동패널과 재정패널 자료를 각각 활용한 연구들의 수가 비교적 크게 차이가 나는 것은 해당 패널조사 자료의 목적과 그에 따른 조사내용이 상이하기 때문이다.

노동패널은 보다 합리적이고 정확한 고용정책의 수립과 실행을 위해 노동시장과 관련 정책의 효과를 분석할 수 있는 기초 자료 제공을 목적으로 한다. 이에 노동패널은 노동시장과 근로자 개인의 근로 경험 등과 관

27) 노동패널 홈페이지에서 제공하는 노동패널 활용목록에서 '소득세', '세제', '조세' 및 '세율' 등의 검색어를 활용하여 관련 논문들을 수집한 결과이다.

28) 2011~2017년 동안의 재정패널학술대회에서 발표된 논문들(81편) 중에서 소득세제에 대한 연구는 18편이고, 이 중에서도 지난 3년 동안 발표된 논문이 13편이다.

련된 세부적인 정보들(교육 및 직업 훈련, 고용상의 특성, 근로시간, 직무만족 및 생활만족, 구직활동, 노동시장 이동 등)을 다른 어떤 패널조사에 비해서도 풍부하게 포함하고 있다. 하지만 근로자 개인들의 소득세액을 계산하기 위한 각종 소득·세액 공제와 관련된 정보는 불충분한 것이 사실이다. 물론, 노동패널에서 제공하는 인적·가구 특성(성, 연령, 경제활동 유무 및 가구주와의 관계 등)과 여러 지출항목(교육, 의료, 주거 및 저축 등)에 대한 정보가 있지만, 소득세제에 적합한 정보를 제공하고 있지는 않다. 예를 들어, 개인연금의 경우에도 세제적격 상품이 아닌 경우 공제혜택을 받을 수 없지만 현재 노동패널에서는 이를 구분할 수 없다. 또한 신용카드 등에 대한 소득공제 금액을 산출하기 위해서는 지출수단(체크카드, 신용카드, 현금)별 지출액을 알아야 하지만, 이에 대한 정보 역시 노동패널에서는 알 수 없다.

반면, 재정패널은 조세정책과 복지정책이 국가와 가계에 미치는 영향과 조세에 대한 부담과 복지 수혜자의 연계성을 분석하고, 조세 모의실험 운용을 위한 데이터베이스를 구축하는 것을 목적으로 한다. 이에 다른 패널자료들과 비교할 때, 재정패널의 가장 큰 특징은 가구원에 대한 조사항목에 신용카드 및 현금영수증, 소득세 납부유형과 소득공제 및 세액공제 현황 등 개인의 납부세액을 계산하기 위한 정보들이 조사·제공된다는 점이다. 이때, “소득공제 내역은 각 항목별로 복잡한 기준과 계산법이 적용되는 경우가 많아 정확한 금액을 회상해내기 어렵기 때문에 소득공제 내역을 증빙할 수 있는 서류를 함께 수집하고 있다. 근로자는 연말정산에 따른 연간 소득내역과 공제내역 및 결정세액 등이 있는 근로소득원천징수영수증을 수집한다. 소득세를 자진 신고하는 종합소득 신고자는 관련 서류를 모두 수집하는 것은 불가능하므로 [과세표준확정신고 및 자진납부계산서]와 [소득공제명세서] 등 2가지 서류를 수집한다.”(송헌재, 2013에서 재인용).

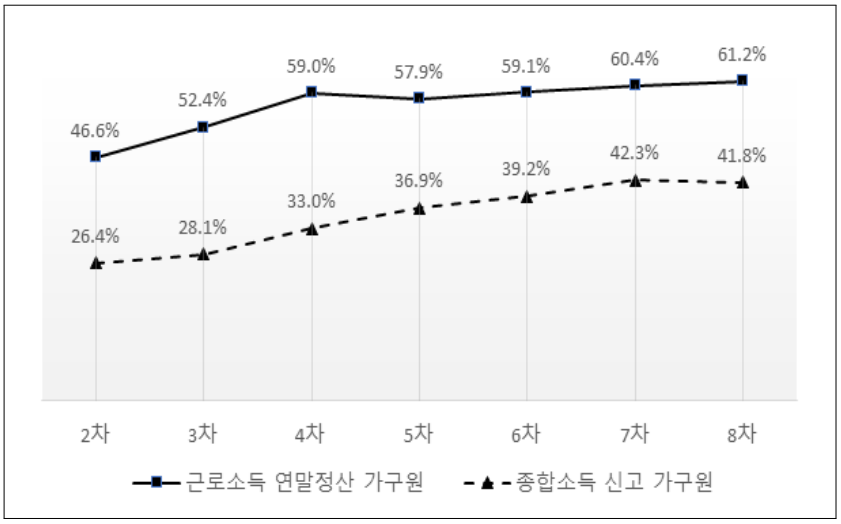
이에 [그림 4-3]을 통해 재정패널 2차에서 8차까지의 소득증빙서류 수집률을 보면[29], 근로소득자의 경우 2차 조사에서 약 46% 수준으로 절반

29) 소득공제증빙서류는 조사대상이 되는 귀속연도의 자료만 수집하는 것이 아니라 과년도의 자료도 지속적으로 수집하고 있다.

에도 미치지 못했지만 이후 점차 증가해 현재는 약 61% 수준이다. 종합소득자의 경우에는 2차 조사 시 약 26%에서 이후 지속적으로 증가하여 최근에는 약 42% 수준이다. 여기서 '수집률'은 근로자의 경우 연말정산을 했다고 응답한 자들, 그리고 종합소득자의 경우 소득신고를 했다고 응답한 자들 중에서 실제로 해당 서류를 수집한 개인들의 비율을 의미한다. 즉 연말정산을 하지 않는 일용 근로소득(자)는 수집대상에서 제외됨을 의미한다.

이처럼 소득증빙서류 수집[30]을 통해 제공되는 개개인의 소득공제 및 세액공제 현황 등에 대한 정보는 개인의 납세액 계산의 정확성 제고에 크

〔그림 4-3〕 소득증빙서류 수집률(2차~8차)



주 : 수집률은 '수집 건수/해당 케이스 숫자'이며, '해당 케이스'는 연말정산을 했다고 응답한 개인을 의미함.
자료 : 한국조세재정연구원, 「각 차수별 기초분석보고서」.

30) 다만, 재정패널의 경우 소득증빙자료 수집률이 좀 더 제고될 필요가 있다고 판단된다. 4차 조사 이후 수집률이 크게 변화가 없다는 점에서 향후 그 비율이 얼마나 증가할 것인지 예측하기가 어려운데, 이러한 상황에서 해당 자료만을 활용한 분석결과가 모집단을 얼마나 대표할 것인지는 의문이다. 특히, 소득증빙서류를 제출하지 않은 개인들이 무작위(random)가 아니고 특정 집단에서 나타난 것이라면 분석결과의 대표성과 신뢰성에 문제가 있을 수 있다는 점은 신중하게 접근할 필요가 있다.

게 기여할 것으로 판단된다. 또한 재정패널의 이와 같은 특징이 해당 자료가 조사·제공된 이후부터 소득세제에 대해 관심을 가진 대부분의 연구자들이 다른 패널자료보다는 재정패널을 활용하는 주요한 이유라고 판단할 수 있다. 왜냐하면 다른 패널자료를 사용하여 근로자 개인의 소득세액을 산출해야 하는 경우 발생하는 임의적인 가정에 따른 편의와 기술적 어려움에 따른 비용을 재정패널을 사용할 때 크게 줄일 수 있기 때문이다.

하지만 재정패널을 활용한다고 해서 모든 어려움이 사라지는 것은 아니다. 특히, 재정패널의 경우 소득공제 항목들 중에서 인적공제에 대한 내용은 제공하지 않고 있는데, 인적공제액 산출 역시 쉽지 않다고 할 수 있다. 왜냐하면 현실에서는 많은 납세자가 함께 살고 있지 않는 부모 혹은 자녀들에 대해 인적공제를 받는 경우(소위 원격지 공제)가 많은데, 조사자료의 정보만으로는 이를 충분히 반영하기가 어렵기 때문이다. 또한 한 가구에 여러 세대가 함께 사는 경우 각 세대별로 인적공제를 분리·할당하는 것이 기술적으로 어려울 수 있다. 예를 들어, 조사대상 가구들 중에는 조손가구 혹은 조카 등 세제상 기본공제대상자가 아닌 아동이 한 가구에 동거하는 경우를 어렵지 않게 확인할 수 있다. 이때, 관련 연구들에서 보면 가구 내 최고 소득자에게 대부분의 공제를 적용하는데, 이것은 가구 내 최고소득자의 공제액을 과대 계산하는 편의와 소득이 있는 다른 가구원의 공제액을 과소 계산하는 편의를 동시에 발생시킨다. 따라서 이러한 편의를 줄이기 위해선 가구원들의 가구 내 관계와 가구변화(분가·합가) 등에 대한 정보를 가공하여 가구원들의 피부양관계를 파악할 필요가 있다. 하지만 이러한 가공과정이 기술적으로 복잡하다는 점에서 쉽지 않은 것이 사실이다. 더욱이 전체 소득공제액 중에서 인적공제액이 차지하는 비중이 시기에 따라 조금씩 다르기는 하지만 약 40% 내외로 가장 큰 비중을 차지한다는 점에서 주의가 필요하다.[31]

정리하면, 소득세추정모형 개발과 관련하여 재정패널이 갖는 장점이 분명히 존재하기는 하지만, 그렇다고 해서 해당 모형의 개발과 활용상의

---

31) 「국세통계연보」를 통해서 전체 소득공제액에서 인적공제액이 차지하는 비중을 보면, 2009~2013년 기간에는 37~39% 수준이다. 소득공제항목이 세액공제로 전환된 이후, 즉 2014~2015년에는 그 비율이 약 46%에 이른다.

모든 어려움이 사라지는 것은 아니다. 이는 재정패널을 활용하여 소득세제의 경제·사회적 영향을 분석하는 연구들이 상대적으로 더 많지만, 구체적인 방법론 측면에서 보면 노동패널 등 다른 패널자료를 활용한 연구들과 큰 차이가 없다는 점에서 다시 한 번 확인할 수 있다.

이러한 맥락에서 소득세제 정책과 관련하여 노동패널 혹은 재정패널을 활용한 연구들의 특징들을 살펴보고자 한다.

먼저, 노동패널을 활용한 연구들은 앞서 말한 바와 같이 6편에 지나지 않지만, 그 주제들을 보면 세제 정책이 개인연금과 저축 혹은 노동공급에 미치는 영향, 그리고 세제개혁의 재분배 및 빈곤효과로서 이하에서 살펴볼 재정패널을 활용한 연구들과 크게 다르지 않다.[32] 특히, 김민성·오시진(2010)에서는 세수중립 제약하에서 초과누진세율을 단일세율로, 그리고 다양한 소득공제를 정액보조로 각각 대체하는 세제 개편이 개인의 세후소득에 미치는 영향을 고찰하는 과정에서 정태적 소득세추정모형을 개발·활용하고 있다. 즉 노동패널에서 제공하는 정보에 최대한 기반하여 당시 세제상의 소득공제액들을 산출하여 개인의 납세액을 계산하고 있다. 나아가 세제개편(단일세율, 정액보조)안에 따른 개인의 납세액을 계산하여, 세제개편이 세후소득 분포에 미치는 영향은 물론, 제도개혁에 따른 소득계층별 편익과 비용을 검토하였다. 다만, 소득세추정모형 개발과 관련된 구체적인 과정과 그 결과에 대한 검토 내용이 상세하게 서술되어 있지 않은 점은 아쉬움으로 남는다.

다음으로 전승훈 외 2인(2008)은 개인연금에 대한 소득공제 혜택의 변화가 가처분소득을 변화시킬 것이고 이로 인한 소득효과가 개인연금 불입액을 얼마나 증가시키는가에 대해 분석한다. 따라서 동 연구에서는 세부담의 변화를 측정하는 것이 매우 중요한데, 저자들은 국세통계연보에서 제공하는 집계값(소득분위별 종합소득세 자료)을 활용하여 계산한 소득분위별 유효세율(소득분위별 과세소득 대비 소득세 부담액)을 이용한

32) 노동패널을 활용한 연구들과 달리, 재정패널을 활용한 연구에서만 이루어진 연구주제는 기부금 공제와 신용카드 등에 대한 공제에 대한 것이다. 현실의 세제에서는 기부금의 유형(정치자금, 지정, 법정)과 지불수단(신용카드, 현금, 체크카드, 전통시장)에 따라 공제율과 한도가 상이한데, 재정패널에서는 이에 대한 정보를 상대적으로 더 상세하게 제공하기 때문이다.

다. 또한 소득공제 혜택에 따른 가처분소득의 변화를 동일한 개인을 대상으로 측정하는 것이 아니라, 서로 다른 세제가 적용되는 여러 시점의 자료에서 소득분위별 개인연금 불입액의 차이를 통해 세제혜택의 효과를 분석한다. 정리하면, 동 연구에서 소득공제혜택에 따른 가처분소득의 탄력성과 세부담 변화에 대한 측정이 매우 중요한데, 이를 집계자료와 나름의 가정하에서 추정한다. 하지만, 해당 지표들을 소득세추정모형으로 측정했다면 좀 더 유효한 결과를 얻을 수 있었을 것으로 판단된다.

김병권 외 3인(2013)에서는 2006년 개정된 연금저축 소득공제 한도 상향이 가계의 저축행태에 미치는 영향에 대해 이중차이분석(DID)을 수행하고 있다. 이를 위해 2005년 개인연금 납부규모가 연간 300만 원보다 적은 가구가 세제혜택 증가(연금저축 소득공제 한도 상향)로 인해 납부금을 늘릴 유인이 있는 가구라는 점에서 처리집단(treatment group), 그리고 300만 원 이상 납부하였던 가구를 통제집단(control group)으로 정의한다. 이를 통해 기존 개인연금 납부규모가 적었던 가구의 개인연금 납부규모가 증가했지만 개연연금 납부 증가로 인해 가구의 다른 저축을 구축한다는 결과를 제시하고 있다. 한편, 동 연구에서 소득세추정모형을 통해 세제혜택에 따른 저축의 탄력성을 추정·활용할 수 있었다면, 정책적으로 좀 더 유용한 결과를 얻을 수 있었을 것이다.

이러한 이중차이분석은 별도의 소득세추정모형이 존재하지 않는 경우 세제개혁의 효과를 분석하는 데 많이 활용되고 있는 분석모형이다. 예를 들어, 세율인하 혹은 근로장려세제가 노동공급에 미치는 영향을 분석하고 있는 심욱기(2006), 남재량(2017), 홍민철 외 2인(2016)에서도 이중차이분석(DID)을 수행하고 있다. 또한 이하에서 살펴볼 재정패널을 활용한 많은 연구 역시 이중차이분석을 수행하고 있다. 이러한 이중차이분석에서는 분석대상을 실험·통제 집단으로 어떻게 구분하는가가 매우 중요하다. 특히, 세제개혁에 따라 그룹 구성원이 바뀌게 될 수 있는데, 이를 통제하기 위한 도구변수의 사용 등 분석에서 고려해야 할 문제들에 대해 상당한 주의를 기울일 필요가 있다. 왜냐하면 홍민철 외 2인(2016)에서는 근로장려세제가 노동공급에 미치는 영향을 분석하는 기존 실증연구들의 결과가 상반되어 나타나는 주요한 원인으로 각 연구마다 서로 다르게 정

의하고 있는 처리·통제 집단의 차이를 지적하고 있기 때문이다.

또한 이중차이분석만으로는 세제개편에 영향을 받는 집단과 그렇지 않은 집단이 명확히 구분될 것을 전제로 하여 그 차이만을 비교·파악할 수 있다는 점에서 한계가 있다. 만일 세제개편으로 인한 영향이 모든 개인에게 나타나는 상황에서 특정 집단에게 그 영향이 상대적으로 크게 나타난다고 할 때, 이중차이분석만으로는 모든 개인에게 나타나는 공통적인 영향은 파악할 수 없을 것이다. 따라서 정책적으로 좀 더 유용한 시사점을 얻기 위해서는 세제개편으로 인해 관련된 모든 개인의 행태 변화와 그 정도를 분석할 수 있어야 할 것이다. 다시 말하면, 소득세추정모형을 활용한다면 세제개편으로 인해 영향을 받는 개인들의 규모와 범위를 확인할 수 있는 동시에, 각 개인의 한계유효세율의 변화를 측정할 수 있을 것이다. 바로 이 점이 소득세추정모형을 개발하기 위한 주요한 이유 중의 하나라고 평가할 수 있다.

다음으로 재정패널을 활용한 연구들(18개 논문)의 내용을 보면, 개인연금 공제가 저축에 대한 유인, 특히 소득계층별 저축에 미치는 효과에 대한 연구가 7편으로 가장 많다. 다음으로 최근 소득공제의 세액공제로의 전환과 관련하여 소득세 체계가 분배상태와 세수, 그리고 근로유인 등의 개인행태에 미치는 영향에 대한 연구가 5편, 기부금 공제에 대한 연구가 4편, 그리고 신용카드와 근로소득공제에 대한 연구가 각각 1편씩 있다. 이에 각 연구들의 주요 내용과 특징을 살펴보면 다음과 같다.

먼저, 박종선·정세은(2017)과 김성태·임병인(2013)의 연구는 소득증빙서류의 정보 외에도 재정패널 자료에서 제공하는 개인·가구에 대한 다양한 정보를 이용하여 임금근로자 개개인의 납세액을 추계한다는 점에서 정태적 소득세모형 개발과 활용 차원에서 볼 때 가장 주목할 만한 연구이다. 박종선·정세은(2017)은 지난 10년간 조세정책의 재분배효과와 세부담의 형평성 분석을 위해 9차 자료에 기반하여 2008년, 2013년, 그리고 2014년 소득세제를 적용하고 있다. 즉 국가 간 서로 다른 제도가 아닌, 동일 국가의 시점별 '정책 스와핑' 분석을 통해 전체 소득체계와 각종의 소득공제항목에 따른 재분배효과와 형평성에 대한 포괄적인 분석을 시도하고 있다. 김성태·임병인(2013) 역시 2013년 당시 세제하에서 각종 공

제항목들을 적용하여 개인의 세부담을 추계하는 한편, 당시 발표되었던 소득공제의 세액공제로의 전환에 따른 소득세제의 형평성 변화를 분석하고, 각각의 공제항목들이 소득분포에 미치는 영향을 포괄적으로 검토하였다.

한편, 송헌재・신유리(2017)과 전병욱(2016)은 소득공제의 세액공제로의 전환이 개인의 근로유인 및 관련 지출(기부, 교육, 의료)에 미친 영향을 분석하고 있다. 송헌재・신유리(2017)는 원천징수자료를 제출한 자들만 분석대상으로 하여 소득공제의 세액공제로의 전환 이전인 2013년과 이후 시점인 2014~15년의 노동공급 변화를 추정・분석하였다. 이때, 정책에 영향을 받는 집단(처리집단)과 그렇지 않은 집단(비교집단)을 과표의 변화 여부를 기준으로 구분하였다. 그 결과, 소득수준에 따른 세제개혁의 영향을 확인하기 힘든 반면, 성별에 따라 그 영향(남성의 경우 근로시간 증가, 여성의 경우 변화 없음)이 다르다는 연구결과를 제시한다. 즉 세제변화에 따른 노동공급 탄력성을 추정・비교하지 않고, 비교집단 대비 처리집단의 상대적인 변화만을 분석하고 있다는 점은 한계라고 평가할 수 있다.

다음으로 전병욱(2016)은 보장성보험료・교육비・의료비・기부금 특별공제가 세액공제로 전환됨에 따라 소득계층별로 해당 부문에 대한 지출변화를 비교・분석하고 있다. 일반적으로 가구별 한계세율이 고소득계층일수록 높다는 점에서 세액공제로의 전환에 따라 고소득계층의 해당 부문들에 대한 지출이 감소하는 반면, 저소득계층의 경우에는 증가할 것으로 기대할 수 있다. 하지만 분석결과 해당 부문에 대한 지출은 소득계층과는 무관한 것으로 보이며, 위험회피적이고 기본적인 지출(교육과 의료)은 비경제적인 요인에 의존하며, 개인적 신념에 의존할 것으로 간주되는 기부금은 경제적 요인에 영향을 받는 것으로 나타난다(전병욱, 2016). 나아가 전병욱(2017)에서는 세액공제로의 전환에 따른 기부금의 절세효과의 변화가 유형별 기부금에 미친 영향을 분석하였다. 이를 통해 가구별 종교에 따른 종교적 신념이 기부금에 영향을 주기도 하지만, 전반적으로 볼 때 기부금의 절세효과, 즉 경제적 유인이 미치는 영향 역시 중요하다는 결과를 제시하고 있다. 한편, 이러한 연구결과는 한도초과기부액 제도

가 가구별 기부금에 미치는 영향에 대한 동 저자의 기존 연구(전병욱・선우희연・정운오, 2015) 결과와도 유사하다. 반면에 박명호(2016)에서는 기부금의 세액공제 전환이 가구의 기부금에 미치는 영향이 부정적이지는 않은 것으로 나타나고 있다. 동 연구에서는 세제변화에 의해 영향을 받는 집단(고소득계층과 저소득계층)과 그렇지 않은 집단(중소득계층)을 과표를 기준으로 구분하여 세제 변화 전후의 집단 간 기부행위의 차이를 DID 분석을 수행한 것이다.

기부금에 대한 이상의 연구들에서 공통적인 점은 제도변화에 따른 기부금의 절세액을 구체적으로 추계한 것은 아니며, 구체적인 수준은 다르지만 가구가 속한 과표와 그에 따른 한계세율을 기준으로 분석대상 가구를 식별하여 분석했다는 점이다. 이와는 달리, 송헌재(2013)는 기부금 공제의 세액공제 전환에 따른 기부금의 가격탄력성을 추정하고 있다는 점에 주목할 만하다. 다만, 좀 더 정확한 추정을 위해 원천징수영수증을 제출한 근로소득자에 대해서만 분석하였다는 점은 아쉬운 점이다. 여하튼 송헌재(2013)는 기부금의 가격탄력성이 7.5~9.9로 상당히 높게 나타남으로써 세제변화에 따른 세수증가효과보다 민간 기부가 축소되는 효과가 더욱 크게 나타날 것으로 전망하고 있다. 다시 말하면, 기부금의 가격탄력성을 추정하여 기부금 관련 조세정책을 검토한다면, 그렇지 않은 경우에 비해 세수변화 혹은 좀 더 세밀한 개혁방안 등 좀 더 정책적으로 유용한 다양한 결과들이 산출될 수 있을 것으로 판단된다.

지금과 같은 방법론 차원에서 아쉬운 점은, 마지막으로 살펴볼 개인연금에 대해 세제혜택 변화가 미치는 영향에 대한 일련의 연구들(김대환, 2017; 정원석・강성호, 2013, 2015, 2016; 김재호, 2015; 강성호, 2014)에 있어서도 마찬가지라는 것이다. 사실, 공적연금의 축소로 노후준비 차원에서 사적연금의 역할 확대는 매우 중요한 정책과제여서 이에 사적연금 가입 확대를 위한 세제개혁이 매년 이루어지다시피 하였다. 이러한 배경하에 상대적으로 많은 연구가 이루어졌다고 할 수 있는데, 이들 연구는 주로 해당 세제개혁이 소득계층별 사적연금 가입에 미치는 영향을 검토하고 있다. 이때, 이들 연구에서 주로 활용된 방법 역시 정책변화 이전과 이후의 저소득층과 고소득층 간 기여액의 차이 변화에 대한 DID 분석방법이

다. 이를 통해 최근 사적연금에 대한 세제혜택이 사적연금, 특히 개인연금 기여액을 전반적으로 증가시키기는 했지만, 이러한 확대는 주로 고소득층을 중심으로 이루어졌음을 보여주고 있다. 즉 노후준비의 필요성이 상대적으로 큰 저소득계층은 오히려 개인연금 가입과 기여가 축소되었다는 것은 해당 정책의 형평성에 심각한 문제일 수 있다. 이처럼 해당 연구결과들이 정책적으로 매우 유용한 것은 사실이지만, 좀 더 구체적인 정책마련을 위해선 그동안의 세제혜택의 분포, 혹은 세제변화에 따른 개인연금의 가격탄력성 추정 등에 대한 분석결과들이 필요할 것으로 판단된다.

제3절 소득세 추정모형의 특징

본절에서는 지난 2년 동안 수행된 2007~2014년 기간 동안의 소득세액 추정을 2015년 귀속 소득세액까지 업데이트한 결과를 간략하게 살펴본다.[33] 다음으로 노동패널에서 조사·제공하는 두 변수, 즉 세전연간소득과 세후연간소득 변수의 차이로 산출할 수 있는 세액(이하, 조사세액)과 본 모형을 통해 추정한 세액(이하, 추정세액)의 차이를 비교·검토한다.

1. 2015년 귀속 소득세액 추정 업데이트 결과

<표 4-2>는 근로소득세 납세자 규모와 비중에 대한 추정결과를 국세통계와 비교한 결과이다.[34] 먼저, 과세대상자에 대한 추정결과가 국세통

33) 소득세 추정 과정에 대한 상세한 설명은 홍민기 외 5인(2016)을 참조. 한편, 업데이트 과정에서 발생한 오류들을 수정함으로써 추정결과들이 조금씩 다르게 나타났음을 밝힌다. 특히, 과거 작업에서 가중치 변수를 잘못 적용함으로써 과세대상 인구와 납세자수 규모 추정이 과소하게 이루어졌는데, 이번 작업에서 해당 오류를 수정했음을 밝힌다.

34) 종합소득세에 있어서 『국세통계연보』는 단순경비율에 의한 추계신고자로서 과세미달자와 비사업자로서 과세미달자는 제외한 통계만을 제공함으로써, 추정결과와 비교가 불가능하다.

계에 비해 약 100만~200만 명 정도 높게 나타나고 있다. 다만, 노동패널도 다른 패널조사 자료들과 마찬가지로 고소득 계층에 대한 표본이 상대적으로 과소하다는 점을 고려하면, 저소득 계층에 대한 과대추정 정도가 상대적으로 더 클 것으로 예상된다. 하지만 근로소득금액이 1천만 원 이하 계층의 경우는 국세통계에 비해 추정 규모가 상대적으로 적게 추정된다는 점에 유의할 필요가 있다.

한편, 납세자 비중 역시 국세통계에 비해 추정결과가 높게 나타나고 있다. 다양한 소득공제 항목들을 비롯해 세액 계산에 필요한 정보들이 부족하기 때문에 분석대상 개인들의 세액이 높게 추정될 가능성이 높고, 이는 납세자 비중을 높이는 요인이다. 다만, 제2절에서 살펴본 바와 같이 소득공제 항목들이 세액공제로 전환한 2014년 이후 납세자 비중이 이전과는 달리 국세통계에 비해 낮게 추정된 결과는 향후 좀 더 주의 깊게 살펴볼 필요가 있다.

〈표 4-2〉 근로소득세 납세자 수와 비중

(단위 : 만 명, %)

	국세통계			노동패널		
	과세 대상자	납세자	납세자 비중	과세 대상자	납세자	납세자 비중
2007		775		1,041	692	66.4
2008		798		1,432	955	66.7
2009	1,425	854	59.9	1,544	1,002	64.9
2010	1,514	924	61.0	1,684	1,145	68.0
2011	1,548	993	64.2	1,709	1,178	68.9
2012	1,572	1,061	67.5	1,756	1,237	70.4
2013	1,630	1,124	69.0	1,814	1,333	73.5
2014	1,663	866	52.1	1,773	870	49.1
2015	1,726	923	53.5	1,838	925	50.3

주 : 1) 납세자는 “소득세액>0”인 자임.
2) ‘노동패널(가중치)’은 개인의 횡단면 가중치를 소수점 아래 첫째자리에서 반올림하여 적용한 결과이며, 2008년 이후부터는 통합가중치를 사용한 것임.
3) 국세통계의 경우 2007~08년에는 과세미달자를 보고하지 않음.

<표 4-3>은 근로소득금액 구간별 납세자수와 평균 세액에 대한 추정치를 국세통계와 비교한 결과이다. 소득금액 구간별 납세자수에 대한 추정결과를 전반적으로 보면, 저소득계층의 납세자수는 국세통계에 비해 많게 추정되는 반면, 소득수준이 높아질수록 납세자수 규모에 대한 추정

〈표 4-3〉 근로소득금액 구간별 평균 결정세액 : 최근 3년

(단위 : 만 원, 만 명, %)

	근로소득금액 구간	납세자수(비중)		평균 세액	
		국세통계	추정치	국세통계	추정치
2013	1천만 원 이하	291 (39.0)	330 (47.2)	6.0	6.9
	2천만 원 이하	254 (89.3)	425 (82.2)	22.6	22.6
	4천만 원 이하	295 (97.4)	390 (95.0)	90.8	88.9
	6천만 원 이하	149 (99.8)	135 (100.0)	273.5	277.4
	8천만 원 이하	62 (100.0)	31 (100.0)	554.3	575.7
	1억 원 이하	29 (100.0)	15 (100.0)	937.2	926.0
	2억 원 이하	23 (100.0)	8 (100.0)	1,941.0	1,756.5
2014	1천만 원 이하	92 (13.8)	57 (10.8)	3.2	2.7
	2천만 원 이하	223 (63.3)	261 (50.1)	15.9	13.6
	4천만 원 이하	262 (74.3)	343 (66.8)	96.0	92.0
	6천만 원 이하	158 (96.9)	139 (96.3)	273.3	270.5
	8천만 원 이하	68 (99.7)	36 (100.0)	600.5	633.4
	1억 원 이하	32 (99.8)	24 (100.0)	1,030.7	1,086.1
	2억 원 이하	27 (99.6)	8 (100.0)	2,099.9	2,265.7
2015	1천만 원 이하	91 (13.5)	51 (10.2)	3.1	2.4
	2천만 원 이하	239 (65.0)	276 (51.0)	16.0	13.3
	4천만 원 이하	277 (76.3)	377 (66.4)	96.7	97.5
	6천만 원 이하	165 (97.4)	138 (96.6)	278.4	275.0
	8천만 원 이하	80 (99.8)	52 (100.0)	596.8	595.3
	1억 원 이하	35 (99.8)	21 (100.0)	1,033.3	1,073.3
	2억 원 이하	31 (99.7)	9 (100.0)	2,105.6	2,292.5

주 : 1) 납세자는 "소득세액>0"인 자임.
2) '노동패널(가중치)'은 개인의 횡단면 가중치를 소수점 아래 첫째자리에서 반올림하여 적용한 결과임.

〈표 4-4〉 종합소득금액 구간별 평균 결정세액 : 최근 3년

(단위 : 만 원, 만 명)

	종합소득금액구간	납세자수		평균 세액	
		국세통계	추정치	국세통계	추정치
2013	1천만 원 이하	127	141	14.0	15.6
	2천만 원 이하	89	169	43.6	60.1
	4천만 원 이하	67	234	161.8	210.7
	6천만 원 이하	30	108	373.0	483.2
	8천만 원 이하	16	43	677.0	801.7
	1억 원 이하	10	18	1,063.9	1,358.1
	2억 원 이하	15	33	2,307.8	2,782.4
2014	1천만 원 이하	170	83	4.8	14.7
	2천만 원 이하	91	159	36.8	53.8
	4천만 원 이하	73	247	154.5	209.9
	6천만 원 이하	29	105	391.3	489.8
	8천만 원 이하	16	48	729.6	839.6
	1억 원 이하	10	20	1,137.9	1,420.1
	2억 원 이하	16	36	2,466.0	2,881.1
2015	1천만 원 이하	197	90	4.4	16.3
	2천만 원 이하	102	158	37.6	58.7
	4천만 원 이하	81	250	154.9	222.8
	6천만 원 이하	32	111	395.1	492.2
	8천만 원 이하	18	58	730.4	873.4
	1억 원 이하	11	21	1,141.4	1,408.0
	2억 원 이하	18	34	2,463.9	2,949.8

주 : 1) 납세자는 "소득세액>0"인 자임.
2) '노동패널(가중치)'은 개인의 횡단면 가중치를 소수점 아래 첫째자리에서 반올림하여 적용한 결과임.

치는 실적치에 비해 상대적으로 적은 것으로 나타나고 있다. 다만, 소득금액이 1천만 원 이하 계층의 납세자 규모와 비중은 2014년 이전과 이후가 다소 상반된 결과가 나타난다. 즉 2014년 이전에는 국세통계에 비해 납세자 비중과 규모가 크게 추정된 반면, 이후에는 상대적으로 작게 추정

되고 있다.

2015년의 경우 평균 근로세액에 대한 추정 결과를 보면, 2014년 추정결과와 유사하게 나타난다. 이는 2015년 당시 세제개혁 폭이 크지 않았던 점이 반영된 것으로 판단된다. 즉 2014년 이후 소득계층별 세부담을 소득공제에서 세액공제로 전환함에 따라 2013년과 비교해 보면, 4천만 원 이하 계층은 줄어들었으며, 4천만~6천만 원 소득 계층은 별다른 변화가 없고, 6천만 원 이상 계층의 세부담은 증가했다.

<표 4-4>는 종합소득세액에 대한 추정결과와 국세통계를 비교한 것이다. 종합소득세 추정치에 대한 비교 결과 역시 2014년과 2015년이 유사하고, 2013년과 비교할 때 2014년 이후 세부담은 저소득계층은 축소, 고소득계층은 증가라는 패턴을 확인할 수 있다. 하지만, 1천만 원 이하 계층에 대한 평균 세액 추정 결과를 보면, 2014년 이후부터는 국세통계에 비해 크게 추정된다는 점에 유의할 필요가 있다.[35]

2. 조사세액과 추정세액 비교

조사세액과 추정세액을 비교하기에 앞서 유의할 점이 있다. 노동패널조사는 조사대상 개인의 '작년 한해 세전 연간 근로소득'과 세후 연간 근로소득을 조사하고 있지만, 이때 작년 한 해 근로소득이 있다고 응답한 모든 자들이 소득세액 추정대상은 아니다. 왜냐하면 소득세액 추정 대상은 직업력 자료의 일자리 정보(종사상지위와 근로기간 등)를 활용하여 개인의 과세유형(근로소득자, 종합소득자)을 정하고 세액을 산출하기 때문이다.[36] 따라서 조사세액과 추정세액에 대한 비교는 소득세액 추정대상이면서 동시에 작년 한 해 근로소득이 있다고 응답한 자들에 대해서만 이루어질 수밖에 없다.

<표 4-5>에서 보면 조사세액과 추정세액 비교대상 인원과 비중은 매

35) 종합소득세에 대한 『국세통계연보』는 단순경비율에 의한 추계신고자로서 과세미달자와 비사업자로서 과세미달자는 제외한 자료라는 점에서 <표 4-4>의 결과는 세액계산 결과 종합소득세액이 '0'인 자를 제외한 결과이다. 하지만, 이들을 포함하는 경우에는 국세통계와 유사한 수준이라는 점에서 향후 검토가 필요하다.

36) 이에 대한 자세한 설명은 홍민기 외 5인(2016)의 제5장 참조.

〈표 4-5〉 비교대상자 선정

(단위 : 명, %)

	총인원	세액추정대상이고 작년소득 있음	세액추정대상이지만 작년소득 없음	세액추정대상 아닌데 작년소득 있음
2007	7,022	5,548 (79.0)	735 (10.5)	739 (10.5)
2008	8,784	6,897 (78.5)	937 (10.7)	950 (10.8)
2009	8,649	7,075 (81.8)	994 (11.5)	580 (6.7)
2010	8,746	7,145 (81.7)	1,025 (11.7)	576 (6.6)
2011	8,893	7,266 (81.7)	1,039 (11.7)	588 (6.6)
2012	8,909	7,309 (82.0)	1,049 (11.8)	551 (6.2)
2013	8,889	7,110 (80.0)	1,253 (14.1)	526 (5.9)
2014	8,999	7,398 (82.2)	1,025 (11.4)	576 (6.4)
2015	9,056	7,501 (82.8)	1,030 (11.4)	525 (5.8)

년 7천 명 내외, 비중은 약 80% 수준임을 알 수 있다. 한편, “세액추정대상 아닌데 작년소득 있음”은 대부분 일용근로자들이다. “세액추정대상이지만 작년소득 없음”의 경우는 약 10% 규모인데, 이들은 향후에 좀 더 신중하게 검토될 필요가 있을 것이다. 왜냐하면 이것은 직업력 자료와 개인조사 자료의 불일치 가능성이 존재하는 것으로서 어떤 정보가 더 신뢰할 수 있는 것인지는 좀 더 면밀한 검토가 필요할 것으로 보인다.

<표 4-6>은 비교대상자들의 납세비중, 즉 조사 혹은 추정 세액이 ‘0’ 이상인 자들의 비중을 비교한 결과이다. 먼저, 근로소득자에 대한 비교 결과를 보면 4천만 원 이상 구간은 대부분 납세자 비중이 100%에 가깝기 때문에 비교할 필요가 없다. 하지만 4천만 원 이하 계층의 경우에는 조사 결과와 추정 결과 간 차이가 상당히 크다. 특히, 1천만 원 이하 구간에서 조사자료상 납세자 비중은 추정결과에 비해 상당히 높게 나타난다. 또한 국세통계상의 납세자 비중을 기준으로 보면 추정치가 국세통계 실적치와 유사한 것을 볼 때, 조사세액(세전소득-세후소득)에 기반한 납세자 비중이 지나치게 높다고 평가할 수 있다.

한편, 종합소득자의 경우는 근로소득자와는 정반대의 결과가 나타난다. 즉 추정결과에 비해 조사자료상의 결과가 상당히 낮게 나타나고 있다.

〈표 4-6〉 소득금액 구간별 납세 비중 비교 : 추정 vs. 조사

(단위 : %)

	소득금액 구간	근로소득세			종합소득세	
		국세통계	추정	조사	추정	조사
2007	1천만 원 이하		44.1	58.0	63.4	14.5
	2천만 원 이하		83.9	90.4	95.8	39.1
	4천만 원 이하		96.9	96.8	99.8	62.6
	6천만 원 이하		100.0	100.0	100.0	73.6
	8천만 원 이하		100.0	97.7	100.0	91.7
	1억 원 이하		100.0	100.0	100.0	92.8
	2억 원 이하		100.0	100.0	100.0	91.4
2008	1천만 원 이하		42.9	63.3	64.7	21.5
	2천만 원 이하		83.5	90.8	95.9	43.0
	4천만 원 이하		96.2	97.8	99.8	64.8
	6천만 원 이하		100.0	99.6	100.0	79.1
	8천만 원 이하		100.0	100.0	100.0	93.4
	1억 원 이하		100.0	100.0	100.0	84.1
	2억 원 이하		100.0	100.0	100.0	86.6
2009	1천만 원 이하	32.4	44.0	60.6	57.9	18.4
	2천만 원 이하	83.9	75.8	91.6	93.3	43.0
	4천만 원 이하	95.2	92.6	95.8	99.5	61.2
	6천만 원 이하	99.7	100.0	99.5	100.0	76.9
	8천만 원 이하	100.0	100.0	100.0	100.0	92.6
	1억 원 이하	100.0	100.0	100.0	100.0	93.4
	2억 원 이하	100.0	100.0	100.0	100.0	81.4
2010	1천만 원 이하	32.7	47.9	61.1	55.2	20.3
	2천만 원 이하	86.8	74.8	91.7	96.0	35.2
	4천만 원 이하	96.4	93.4	97.4	99.5	67.5
	6천만 원 이하	99.8	100.0	99.4	100.0	78.9
	8천만 원 이하	100.0	100.0	100.0	100.0	90.1
	1억 원 이하	100.0	100.0	95.6	100.0	85.9
	2억 원 이하	100.0	100.0	100.0	100.0	87.4
2011	1천만 원 이하	35.3	44.5	63.4	57.5	17.4
	2천만 원 이하	87.3	79.1	91.9	93.3	38.2
	4천만 원 이하	96.4	92.9	95.4	99.9	63.4

〈표 4-6〉의 계속

	소득금액 구간	근로소득세			종합소득세	
		국세통계	추정	조사	추정	조사
2011	6천만 원 이하	99.7	100.0	98.8	100.0	77.2
	8천만 원 이하	100.0	100.0	100	100.0	89.8
	1억 원 이하	100.0	100.0	98.6	100.0	81.0
	2억 원 이하	100.0	100.0	94.4	100.0	93.8
2012	1천만 원 이하	38.0	43.3	61.9	62.8	22.1
	2천만 원 이하	88.4	82.1	90.3	95.0	35.1
	4천만 원 이하	97.0	93.6	97.0	99.8	60.3
	6천만 원 이하	99.8	100.0	99.2	100.0	77.8
	8천만 원 이하	100.0	100.0	100.0	100.0	91.4
	1억 원 이하	100.0	100.0	100.0	100.0	91.2
	2억 원 이하	100.0	100.0	100.0	100.0	91.2
2013	1천만 원 이하	39.0	47.3	60.3	58.0	24.5
	2천만 원 이하	89.3	82.0	89.1	94.6	37.6
	4천만 원 이하	97.4	95.0	96.1	99.5	63.3
	6천만 원 이하	99.8	100.0	99.8	100.0	76.5
	8천만 원 이하	100.0	100.0	100.0	100.0	87.6
	1억 원 이하	100.0	100.0	99.0	100.0	96.4
	2억 원 이하	100.0	100.0	100.0	100.0	86.6
2014	1천만 원 이하	13.8	10.9	58.4	43.3	19.1
	2천만 원 이하	63.3	50.0	85.8	90.5	43.1
	4천만 원 이하	74.3	66.8	94.7	98.6	62.7
	6천만 원 이하	96.9	96.3	99.0	100.0	71.4
	8천만 원 이하	99.7	100.0	100.0	100.0	87.9
	1억 원 이하	99.8	100.0	98.5	100.0	85.2
	2억 원 이하	99.6	100.0	98.2	100.0	87.4
2015	1천만 원 이하	13.5	10.7	61.7	44.6	20.3
	2천만 원 이하	65.0	50.9	86.6	90.2	44.4
	4천만 원 이하	76.3	66.3	95.7	98.1	63.7
	6천만 원 이하	97.4	96.6	99.9	100.0	79.1
	8천만 원 이하	99.8	100.0	100.0	100.0	92.7
	1억 원 이하	99.8	100.0	100.0	100.0	92.2
	2억 원 이하	99.7	100.0	100.0	100.0	92.7

하지만 종합소득세의 납세자 비중에 대한 실적치가 없기 때문에 어떤 통계치가 좀 더 현실적인지를 판단하기 어렵다.

<표 4-7>은 소득금액 구간별로 추정세액과 조사세액을 비교한 결과이다. 평균세액에 대한 비교는 근로소득세와 종합소득세 모두 유사하게 나타난다. 즉 4천만 원 이하 계층에서 조사세액이 지나치게 크게 나타나고 있고, 소득수준이 낮을수록 추정세액과 조사세액의 차이는 더욱 증가한다.

〈표 4-7〉 소득금액 구간별 평균 세액 비교 : 추정 vs. 조사

(단위 : 만 원)

	소득금액구간	근로소득세			종합소득세		
		국세통계	추정	조사	국세통계	추정	조사
2007	1천만 이하	7.8	8.5	57.1	18.2	24.0	36.9
	2천만 이하	32.9	32.4	112.7	65.8	92.7	68.8
	4천만 이하	126.2	126.9	210.2	214.1	283.5	150.2
	6천만 이하	370.9	380.6	412.2	510.2	605.6	264.0
	8천만 이하	735.8	849.2	614.5	912.5	1,086.7	383.9
	1억 이하	1,217.8	1,185.3	1,101.5	1,373.1	1,616.1	517.5
	2억 이하	2,442.9	1,701.1	1,200.0	2,723.3	3,131.6	1,199.8
2008	1천만 이하	7.5	9.0	58.1	17.6	23.5	45.0
	2천만 이하	29.2	27.9	120.6	60.2	75.2	88.6
	4천만 이하	105.0	98.7	227.1	191.8	241.6	169.6
	6천만 이하	327.0	333.7	422.4	461.3	536.5	312.0
	8천만 이하	649.5	683.8	836.1	826.7	914.8	353.3
	1억 이하	1,104.0	1,064.3	1,167.8	1,279.3	1,468.6	653.7
	2억 이하	2,273.4	1,923.8	2,310.2	2,595.8	3,077.4	987.7
2009	1천만 이하	5.4	6.6	54.9	13.1	15.3	39.0
	2천만 이하	20.2	21.7	119.7	42.9	58.0	86.4
	4천만 이하	78.3	84.8	214.5	152.9	213.1	154.6
	6천만 이하	265.0	282.7	395.5	399.1	501.8	265.2
	8천만 이하	555.8	559.1	695.3	735.0	837.4	350.5
	1억 이하	975.3	1,048.3	900.1	1,166.4	1,416.0	407.1
	2억 이하	2,129.1	1,798.8	1,589.4	2,462.3	2,597.7	1,054.2

〈표 4-7〉의 계속

	소득금액구간	근로소득세			종합소득세		
		국세통계	추정	조사	국세통계	추정	조사
2010	1천만 이하	5.6	6.9	52.7	13.3	17.1	37.1
	2천만 이하	21.4	22.2	106.1	43.0	58.5	82.5
	4천만 이하	82.2	81.7	200.5	151.5	207.4	145.2
	6천만 이하	266.7	272.9	405.1	378.2	460.1	281.3
	8천만 이하	560.6	573.5	640.3	698.6	833.8	366.0
	1억 이하	959.9	1,019.6	769.0	1,111.7	1,388.9	419.6
	2억 이하	2,039.4	2,438.2	1,351.2	2,381.4	2,819.5	1,065.8
2011	1천만 이하	5.8	6.4	48.4	13.5	15.3	30.1
	2천만 이하	22.1	21.8	96.4	43.3	59.1	84.1
	4천만 이하	83.8	86.1	198.8	150.6	212.8	133.4
	6천만 이하	262.7	269.4	332.0	386.2	471.2	243.0
	8천만 이하	546.8	595.2	523.5	684.1	844.8	367.7
	1억 이하	933.4	918.8	738.4	1,082.1	1,363.0	493.3
	2억 이하	1,952.5	1,867.7	1,145.7	2,331.4	2,592.9	893.7
2012	1천만 이하	5.9	6.4	49.8	13.8	15.4	42.5
	2천만 이하	22.3	21.5	97.9	43.4	57.4	77.4
	4천만 이하	87.5	83.6	188.9	145.9	207.3	132.2
	6천만 이하	266.8	272.2	359.6	356.0	463.3	248.3
	8천만 이하	548.8	573.6	601.8	665.5	848.9	394.1
	1억 이하	932.9	890.2	693.2	1,064.1	1,322.7	454.3
	2억 이하	1,940.7	1,624.9	1,193.1	2,304.6	2,689.4	820.1
2013	1천만 이하	6.0	6.9	44.0	14.0	16.6	36.2
	2천만 이하	22.6	22.6	88.2	43.6	59.7	64.8
	4천만 이하	90.8	88.9	184.7	161.8	208.0	126.4
	6천만 이하	273.5	277.4	337.7	373.0	479.9	247.3
	8천만 이하	554.3	575.7	509.4	677.0	792.8	299.6
	1억 이하	937.2	926.0	943.2	1,063.9	1,354.9	464.0
	2억 이하	1,941.0	1,756.5	1,274.2	2,307.8	2,763.2	877.6
2014	1천만 이하	3.2	2.7	41.9	4.8	13.5	36.7
	2천만 이하	15.9	13.5	85.7	36.8	53.4	68.3

〈표 4-7〉의 계속

	소득금액구간	근로소득세			종합소득세		
		국세통계	추정	조사	국세통계	추정	조사
2014	4천만 이하	96.0	91.9	184.8	154.5	206.8	142.5
	6천만 이하	273.3	270.5	379.9	391.3	482.0	238.4
	8천만 이하	600.5	633.4	554.2	729.6	837.0	409.4
	1억 이하	1,030.7	1,086.1	952.1	1,137.9	1,411.2	645.8
	2억 이하	2,099.9	2,265.7	1,929.4	2,466.0	2,850.0	1,062.5
2015	1천만 이하	3.1	2.4	46.5	4.4	15.8	45.2
	2천만 이하	16.0	13.3	92.6	37.6	56.6	79.1
	4천만 이하	96.7	97.4	215.1	154.9	222.1	150.0
	6천만 이하	278.4	275.3	385.3	395.1	484.5	363.0
	8천만 이하	596.8	595.3	588.3	730.4	873.4	456.7
	1억 이하	1,033.3	1,073.3	1,038.9	1,141.4	1,389.0	613.9
	2억 이하	2,105.6	2,292.5	1,761.7	2,463.9	2,936.9	1,111.1

앞의 납세자 비중에 대한 결과와 함께 정리하면, 조사자료를 활용하여 납세자 비중과 납세액을 산출하는 경우 지나치게 과대평가할 위험이 있다. 특히, 이러한 과대평가의 위험은 소득수준이 낮을수록 더욱 크다고 평가할 수 있다.

한편, 이와 같은 결과가 나타난 이유는 여러 가지가 있을 수 있다는 점에서 향후 유의할 필요가 있을 것이다. 예를 들어, 조사대상 개인들이 연말정산을 통한 세액환급을 고려하지 않고 응답했을 수도 있다. 혹은 설문조사관에게 자신이 면세자임을 밝히고 싶지 않을 수도 있고, 자신이 납부하는 세금을 실제보다 많게 응답할 가능성도 있다. 또한 세법상 과세대상 소득과 개인들이 응답한 소득의 정의와 범위가 서로 다를 수 있다는 점도 일부 영향을 미쳤을 것이다.

제4절 소 결

본 연구는 소득세제가 소득재분배 및 개인의 경제·사회적 행태에 미치는 영향을 분석하기 위한 미시 모의실험 모형, 즉 소득세추정모형의 필요성과 의의, 그리고 관련 국내외 사례를 살펴보았다. 이것은 노동패널자료에 기반하여 개인들의 납세액을 추정한 지난 2년 동안의 작업을 향후 소득세추정모형으로 발전시키고자 하는 경우 참고할 수 있는 시사점 도출을 목적으로 한다. 이러한 맥락에서 본 연구를 통해 얻을 수 있는 시사점들은 다음과 같다.

첫째, 세제정책의 사전적 평가 혹은 사후적 효과 분석에서 미시자료의 필요성을 다시 한 번 확인할 수 있다. 예를 들어, 세제정책과 관련해서 대부분의 연구들은 세제개혁의 영향이 소득 수준은 물론, 개인·가구의 사회·경제적 특성에 따라 다양하게 나타난다는 점을 지적하고 있다. 즉 동일한 세제개혁일지라도 개개인이 처한 경제·사회적 상황에 따라 어떤 자는 이득을 얻을 수도 있지만 다른 어떤 자는 손실을 입을 수도 있다. 따라서 이를 적절하게 분석하기 위해선 개인 혹은 가구 단위의 다양한 정보를 조사·제공하는 미시자료가 활용될 필요가 있다.

둘째, 소득세제의 정책효과 분석에 있어서 미시자료의 축적이 필요조건이기는 하지만 충분조건은 아니다. 예를 들어, 재정패널에서 제공하는 소득증빙서류에 담겨 있는 소득·세액 공제 등에 대한 정보는 개인의 납세액 계산의 정확성을 크게 높일 수 있는 매우 중요한 정보임에 틀림이 없고, 나아가 이러한 정보를 제공하는 패널자료는 국내는 물론, 국외에서도 찾아보기 힘들다. 하지만 해당 정보가 있다고 할지라도 재정패널에 기반한 소득세추정모형이 지난 약 10년 동안 개발되지 않은 것은 해당 모형의 개발과 관리를 위한 연구와 지원이 반드시 필요한 것임을 보여주는 것으로 이해할 수 있다. 한편, 이러한 맥락에서 아쉬운 점은 성명재 외 2인(2008, 2010)에서 보면 조세재정연구원이 재정패널 자료를 활용하여 조세/복지 모형 개발을 시도한 것으로 보이지만, 이후 해당 모형의 업데이트

및 관리 등의 부재로 후속 연구가 이루어지지 않고 멈춘 것으로 보인다는 점이다. EUROMOD의 사례에서도 본 것처럼, 특정 기관(ISER)이 중심이 되어 모형개발과 유지 및 제도와 자료의 업데이트를 담당하는 한편, 모형의 세부적인 특징과 이를 실행할 수 있는 연구자들에 대한 교육 등을 통해 모형을 활용하는 사용자의 범위를 확대하기 위한 노력이 이루어질 필요가 있다.

셋째, 소득세에 대한 미시 모의실험 모형 분석과 활용의 가장 큰 장점은 좀 더 정교하고 현실적인 정책대안 마련에 있다고 평가할 수 있다. [그림 4-1]과 [그림 4-2]에서 보는 것처럼 소득세 체계는 누진적 세율 외에도 각종 공제항목으로 구성되어 있다. 그 결과, 단순히 누진적 세율구조를 통해서만 세수와 소득재분배 효과가 나타나는 것이 아니라, 다양한 목적의 공제제도들 역시 영향을 미친다. 더욱이 다양한 목적의 공제제도들은 세수와 재분배 외에도 해당 공제항목과 관련된 개인/가구의 행태에도 영향을 미친다는 점에서 소득세 모형의 개발은 세제정책에 상당히 유용할 것으로 판단된다.

넷째, 조세/복지 모형을 활용하는 경우 조세/복지 제도와 개인/가구의 경제·사회적 행태와 관련된 이론적·정책적 연구에 중요한 기초 결과를 산출할 수 있다는 점 역시 주목할 필요가 있다. 세제정책이 개인의 행태에 미치는 영향을 포착할 수 있는 가장 중요한 지표는 한계유효세율(marginal effective tax rate)이다. 왜냐하면 특정의 소득공제가 변화하는 경우 상대적으로 그 범위가 넓은 과표에는 영향을 주지 않을 수 있지만, 실제 납부하는 세액에는 영향을 줄 수 있기 때문이다. 예를 들어, 2016년 기준으로 15%의 세율이 적용되는 과표구간은 1,200만~4,600만 원인데, 이때 특정 공제항목의 금액이 100만 원 정도 변화할 때 기존 과표가 최소 혹은 최대 금액에 인접한 자들 외에는 과표 자체가 달라지지 않을 것이다. 하지만, 과표가 달라지지 않더라도 공제혜택의 증가는 실제 납부하는 세액을 축소시킴으로써 개인의 한계유효세율을 변화시킨다. 다시 말하면, 소득세 체계를 구성하는 다양한 요소들을 반영하여 개인의 납세액을 추계할 때, 세제개혁으로 인한 개개인의 한계유효세율의 변화를 포착가능하다는 점에서 조세/복지 모형 개발과 활용의 필요성과 장점이 있다.

다섯째, EUROMOD의 경우 모형의 신뢰성 제고와 활용범위를 넓히기 위한 다양한 노력을 지속적으로 한다는 점은 향후 관련 모형 개발에 있어서 시사하는 바가 크다. 특히, 기준선 시뮬레이션 결과에 대한 유효성(validation)을 검토하고 이를 공개하고 있다는 점은 주목할 필요가 있다. 사실, 시뮬레이션 분석의 특성상 현실과의 오차는 피할 수 없다. 따라서 모형으로부터 산출된 기본적인 결과를 실적치와 비교·공개함으로써 모형 개발과정에서 있을 수 있는 기술적 오류의 수정·개선, 나아가 모형의 전반적인 발전을 촉진하는 계기로 삼는 것이 타당하다. 또한 EUROMOD의 기술적 특성과 세부적인 방법론 등에 대한 정보 제공과 공유는 새로운 연구자들이 해당 모형을 처음으로 이용할 때 겪게 되는 비용과 노력을 크게 줄여준다. 정리하면, 현재 존재하고 있는 미시 모의실험 모형들 중에서도 EUROMOD가 가장 폭넓게 활용되고 있는 것은 이상의 노력들의 결과라고 평가할 수 있을 것이다. 반면에 국내사례에서 보면 연구자와 연구주제가 사실 상당히 제한적임을 알 수 있는데, 이는 관련 모형의 개발과 공유를 위한 노력의 필요성을 반증하는 것으로 이해할 수 있다.

마지막으로, 지난 2년 동안 노동패널을 활용한 소득세 추정 작업의 결과를 소득세추정모형으로 발전시키고자 한다면, 이를 위한 후속 작업들이 지속적으로 이루어질 필요가 있다. 사실, 지난 2년 동안의 작업은 노동패널 조사 대상 개인들의 소득세액만을 추정하는 것이 애초의 목적이었다. 하지만 이를 다양한 연구에 활용할 수 있도록 많은 연구자와 그 성과를 공유하기 위해선 지금보다 더 많은 노력과 시간이 필요하다.

예를 들어, 소득세액 계산에 필요한 각종 공제항목의 수혜 범위와 규모가 적절하게 이루어졌는지에 대한 검토가 필요하다. 나아가 개선이 필요한 경우 이를 위한 방법론 개발과 같은 모형의 내적인 측면에 대한 연구들도 필요할 것이다. 또한 분석결과의 대표성을 확보하기 위해 고소득자들에 대한 표본을 확충한다거나, 현실의 세제내용을 고려한 조사항목의 추가 등도 고려할 필요가 있다. 그리고 소득세추정모형의 세부적인 방법론상의 특징들을 관심 있는 연구자들과 공유하기 위한 노력과 더불어 좀 더 긴 시간을 필요로 하겠지만 다양한 분야의 연구자들과 정책담당자들이 손쉽게 활용할 수 있는 프로그램 개발과 같은 기술적인 지원도

필요하다.

물론, 이와 같은 노력이 쉽게 이루어진다거나, 혹은 그 성과가 단기간에 나타난다고 감히 보장할 수는 없다. 하지만, 완성된 형태의 소득세추정모형을 개발하고 지속적인 관리를 통해 유지·발전시킨다면, 많은 연구자들과 정책담당자들에게 그 활용범위는 상당히 넓을 것이며, 정책 수립과 실행에 매우 유용한 분석도구가 될 것임은 분명하다. 특히, 노동패널자료가 노동시장과 고용정책 분석에 필요한 정보를 풍부하게 제공하고 있음을 고려하면, 소득세추정모형 개발에 따른 활용가치는 다른 어떤 패널자료에 비해서도 더욱 클 것으로 기대할 수 있다는 점에서 향후에도 많은 관심과 노력이 지속적으로 이루어질 필요가 있다.

제 5 장
한국노동패널과 경제활동인구조사에 나타난 노동시장 활동 비교[37)]

제1절 서 론

개인의 노동시장 활동의 변화를 살펴보는 것은 국가, 지역, 산업의 노동공급과 노동수요의 변화를 측정할 수 있는 가장 기본적인 방법이다. 이러한 이유로 많은 행정기관이나 연구기관에서는 주기적으로 개인의 노동시장 활동에 대한 보고서를 발간하고 있고 이는 정부의 노동시장 정책을 세우는 데 있어서 매우 중요한 역할을 하고 있다.

노동시장 활동을 나타내는 지표로서는 경제활동참가율(=15세 이상 인구 중 경제활동에 참가하는 사람의 비중), 고용률(=15세 이상 인구 중 취업자의 비중), 그리고 실업률(=경제활동인구 중 실업자의 비중) 등이 많이 사용된다. 경제활동참가율은 한 사회의 사용 가능한 노동력을 측정하는 지표로 이용되며, 고용률은 고용상황의 변동을 나타내는 지표로, 실업률은 유휴인력 또는 노동시장의 수요부족을 나타내는 지표로 이용된다.

현재 우리나라에서 개인의 노동시장 활동을 파악할 수 있는 가장 대표적인 자료는 통계청의 경제활동인구조사라 할 수 있다. 경제활동인구조사는 약 33,000가구를 대상으로 하는 가구조사인데 15세 이상의 모든 가

37) 연구조교로 일한 공주대 경제통상학부 이현성, 나현민 학생에게 감사한다.

구원을 대상으로 개인의 특성과 노동시장 활동에 대한 정보를 묻고 있다. 이 조사는 매월 이루어지며 주기적으로 표본이 재설정되고 있다.

경제활동인구조사 외에 개인의 노동시장 활동을 측정할 수 있는 또 다른 자료로는 한국노동연구원에서 제공하는 한국노동패널(KLIPS)이 있다. KLIPS는 1998년 1차 조사를 시작으로 지금 현재까지 18차에 이르고 있다. 또한, 표본이 점점 고령화됨에 따라 2009년에 1,415가구를 새롭게 표본에 추가하여 현재 이 통합표본에 대한 설문조사는 7차에 이르고 있다.

한편, KLIPS는 개인의 노동시장 활동을 추적하는 패널자료(panel data)를 목적으로 하고 있기 때문에 한 시점의 단면(cross-section)을 보는 경제활동인구조사와는 다르지만, 많은 연구자가 KLIPS를 이용하여 반복적 횡단분석(repeated cross-sectional analysis)을 하고 있기 때문에 횡단면 자료로도 많이 사용되고 있다. 또한, KLIPS는 경제활동인구조사와 마찬가지로 인구주택총조사를 모집단으로 하고 있고, 횡단면 가중치를 제공하고 있기 때문에 원칙적으로는 횡단면 자료로서의 대표성도 동시에 보유하고 있다.

본 연구의 목적은 KLIPS의 통합표본이 시작되는 2009년 이후 한국노동연구원에서 공급하고 있는 한국노동패널(KLIPS)에서 측정되는 노동시장 활동 지표(경제활동참가율, 고용률, 실업률)와 통계청의 경제활동인구조사에서 나타나는 지표를 시계열적으로 비교하는 것이다. 또한, 이들의 차이정도를 확인하고, 그 원인과 개선을 위한 제언을 제시하고자 한다.

제2절 연도별 노동시장 활동 현황

한국노동패널(KLIPS)은 1998년부터 시작하였으나 2009년부터 1,415가구를 새롭게 표본으로 추가하여 통합표본을 유지하고 있다. 따라서 시계열의 연속성을 유지하기 위하여 분석기간은 2009년부터 설정하기로 한다. 노동시장 활동 현황으로는 경제활동참가율, 고용률, 그리고 실업률을 중심으로 살펴보며 통계청의 경제활동인구조사와 한국노동연구원의 KLIPS

를 비교하는 것을 목적으로 한다.

1. 전체 근로자

경제활동참가율은 15세 이상 중 경제활동에 참가하고 있는 사람들의 비중인데 경제활동 참가자는 다시 취업자와 실업자로 구분된다. 여기서 주의해야 할 것은 실업자의 기준이 두 가지가 있을 수 있다는 것이다. 첫째는 ILO 기준인데 이는 지난 1주일간 구직활동여부에 따라 결정된다. 둘째는 OECD 기준인데 이는 지난 4주간의 구직활동 여부에 따라 결정된다. 결과적으로 OECD 기준으로 할 경우 실업자의 수가 많아지게 되며 따라서 실업률과 경제활동참가율도 높아지게 된다.

<표 5-1>은 KLIPS와 경제활동인구조사에 나타난 취업자, 실업자, 비경제활동인구와 경제활동참가율, 고용률, 실업률을 OECD 기준으로 구분하여 제시하고 있다. ILO 기준으로 계산된 수치는 <부표 5-1>에 제시되는데 OECD 기준과 큰 차이를 보이지 않기 때문에 OECD 기준을 중심으로 설명하고자 한다.

우선 취업자수를 보면 KLIPS의 취업자가 경활인구조사의 취업자보다 적게 포착되고 있음을 알 수 있다. 2009년, 2010년 KLIPS의 취업자수가 경활조사 취업자수의 94%로 가장 낮게 나타나며 그 이후 98%까지 상승하였다가 2015, 2016년에는 96%에 이르고 있다.

실업자는 취업자보다 괴리가 더 크게 나타나고 있다. 2012년에는 KLIPS의 실업자수가 경활조사의 실업자수의 98%에 달했으나 그 이후 감소하여 2016년에는 그 비중이 67%까지 떨어지고 있다. <부표 5-1>을 보면 ILO 기준의 경우 KLIPS의 실업자가 더욱더 과소하게 포착되어 그나마 OECD 기준으로 할 때 KLIPS의 실업자수가 통계청 수치와 더 근접하고 있다.

반면에 KLIPS는 비경제활동인구를 경활조사보다 과대 포착하고 있는 것으로 나타나고 있다. 경활조사의 비경활인구수 대비 KLIPS의 비경활인구수는 2009년 1.13배에서 2016년 1.11배로 지난 8년간 평균적으로 1.1배 더 높게 나타나고 있다.

〈표 5-1〉 KLIPS와 통계청 기준 경제활동지표(OECD 기준)

		2009	2010	2011	2012	2013	2014	2015	2016
취업자	KLIPS	22,163	22,431	23,799	24,678	24,421	24,893	24,933	25,217
	통계청	23,506	23,829	24,244	24,681	25,066	25,599	25,936	26,235
	KLIPS/통계청	0.94	0.94	0.98	0.98	0.97	0.97	0.96	0.96
실업자	KLIPS	1,051	622	738	806	660	694	683	679
	통계청	889	920	855	820	807	937	976	1,012
	KLIPS/통계청	1.18	0.68	0.86	0.98	0.82	0.74	0.70	0.67
비경활 인구	KLIPS	17,735	17,151	17,471	17,542	17,769	17,638	17,962	18,006
	통계청	15,698	15,841	15,953	16,081	16,223	15,977	16,105	16,169
	KLIPS/통계청	1.13	1.08	1.10	1.09	1.10	1.10	1.12	1.11
경제 활동 참가율	KLIPS	56.7	57.3	58.4	58.7	58.5	59.2	58.8	59.0
	통계청	60.8	61.0	61.1	61.3	61.5	62.4	62.6	62.8
	KLIPS/통계청	0.93	0.94	0.96	0.96	0.95	0.95	0.94	0.94
고용률	KLIPS	54.1	55.8	56.7	56.8	57	57.6	57.2	57.4
	통계청	58.6	58.7	59.1	59.4	59.5	60.2	60.3	60.4
	KLIPS/통계청	0.92	0.95	0.96	0.96	0.96	0.96	0.95	0.95
실업률	KLIPS	4.5	2.7	3	3.2	2.6	2.7	2.7	2.6
	통계청	3.6	3.7	3.4	3.2	3.1	3.5	3.6	3.7
	KLIPS/통계청	1.25	0.73	0.88	1.00	0.84	0.77	0.75	0.71

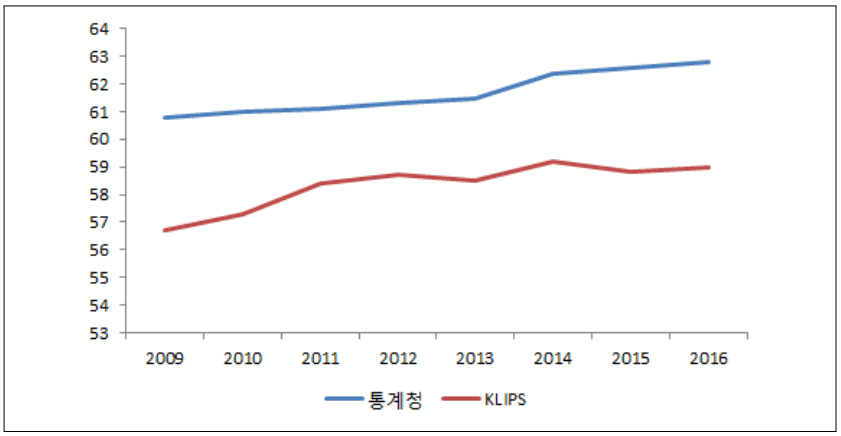
주 : KLIPS는 횡단면 가중치 사용, 통계청은 KOSIS 경제활동인구조사 연자료.

한편, KLIPS에서 계산된 경제활동참가율과 고용률은 경활조사에서 구해진 것의 약 95% 수준을 나타냄으로써 큰 차이를 보이지는 않고 있다. 하지만 실업률은 상대적으로 큰 괴리를 보이고 있다. 2009년을 제외하고는 KLIPS의 실업률도 경활조사의 실업률보다 낮게 나타나는 경향을 보이고 있다. 2016년에는 KLIPS 실업률이 2.6%인 데 반하여 경활조사의 실업률은 3.7%로 약 1.1%p의 차이를 나타내고 있다.

경제활동참가율, 고용률, 실업률의 크기도 중요하지만 이들의 시계열적 변화도 역시 중요하다. 통계청 경활조사의 지표와 KLIPS의 지표 간 차이가 있더라도 그 변화의 경향이 유사하다면 추이를 판단하는 데는 큰 무리가 없을 것이기 때문이다. [그림 5-1]은 경제활동참가율을 보여주고 있는

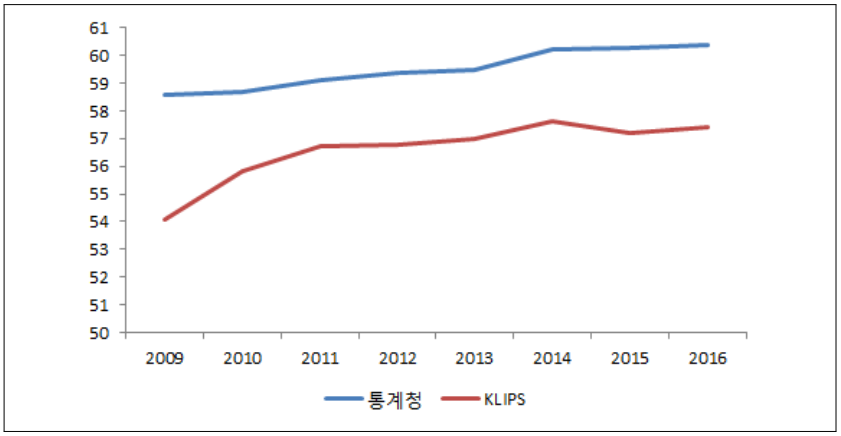
데 통계청의 경제활동참가율이 KLIPS보다 높게 나타나고 있지만 두 자료 사이의 상관계수가 0.775로 서로 양의 상관관계를 보이고 있다. [그림 5-2]는 고용률의 시계열 추이를 보여주는데 이 역시 두 자료 사이에 강한 양의 상관관계를 보이고 있다. 따라서 KILPS를 사용하였을 경우 이 두 지표의 추이는 통계청의 경활자료를 사용하였을 경우와 유사하게 나타난다고 말할 수 있을 것이다.

〔그림 5-1〕 경제활동참가율



주 : 두 자료의 상관계수는 0.775로서 95% 수준에서 통계적으로 유의함.

〔그림 5-2〕 고용률

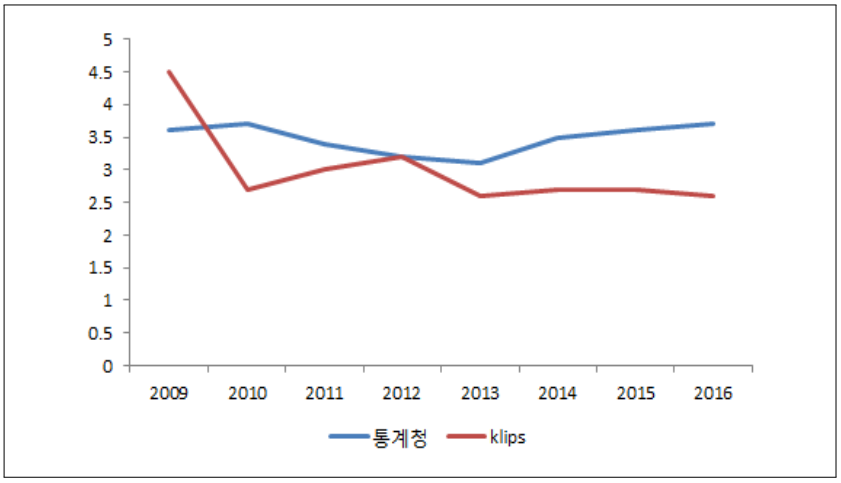


주 : 두 자료의 상관계수는 0.832로서 95% 수준에서 통계적으로 유의함.

한편, [그림 5-3]을 보면 실업률의 경우에는 통계청의 경제활동인구조사를 사용하였을 경우와 KLIPS를 사용하였을 경우 매우 다른 추이를 보인다는 것을 알 수 있다. 또한, 2009년 KLIPS의 실업률이 통계청 실업률보다 과대하게 추정되는 것을 제외하고는 KLIPS의 실업률이 전반적으로 낮게 나타나고 있다. 중요한 것은 추세인데 두 시계열 사이에 상관계수가 0.079로서 90% 수준에서도 통계적으로 유의하지 않는 것으로 나타나고 있다. 다행히 2012년 이후에는 두 시계열 자료가 서로 비슷한 방향으로 움직이는 것으로 나타나고 있다. 하지만 2012년 이전의 경우에는 KLIPS를 이용하여 얻어진 실업률로 통계청 실업률의 변화를 추론하기는 어렵다고 판단된다.

이상을 종합하면 KLIPS는 통계청의 경제활동인구조사와 비교하여 취업자와 실업자를 과소 포착하고, 비경활인구를 과대 포착하고 있다. 하지만 KLIPS의 경제활동참가율과 고용률은 경활조사 해당 수치의 95% 수준에 접근하고 있기 때문에 활용에 큰 문제는 없다고 판단된다. 또한, 시계열적으로 보더라도 KLIPS에서 구해진 경제활동참가율과 고용률은 경활조사에서 얻어진 해당지표와 매우 높은 양의 상관관계를 가진다. 다만 문제가 될 수 있는 것은 실업률이라고 할 수 있다. KLIPS에서는 실업자를

〔그림 5-3〕 실업률



주 : 두 자료의 상관계수는 0.079로서 90% 수준에서 통계적으로 유의하지 않음.

과소 포착하고 있고, 시계열적으로 보더라도 KLIPS와 경활조사 사이에 유의미한 상관관계를 보이지 않고 있다. 때문에 KLIPS를 이용하여 실업자의 노동시장 이행을 분석한다든지 아니면 실업률을 추정하는 것은 대표성 측면에서 주의해야 할 것으로 사료된다.

2. 남성 근로자

남성의 노동시장 활동에 대한 통계는 <표 5-2>에 제시된다. KLIPS에서 포착되는 남성 경제활동인구(취업자와 실업자)는 통계청 경활조사에서 포착되는 남성 경활인구보다 여전히 작지만 그 격차는 전체 근로자에

〈표 5-2〉 남성 KLIPS와 통계청 기준 경제활동지표(OECD 기준)

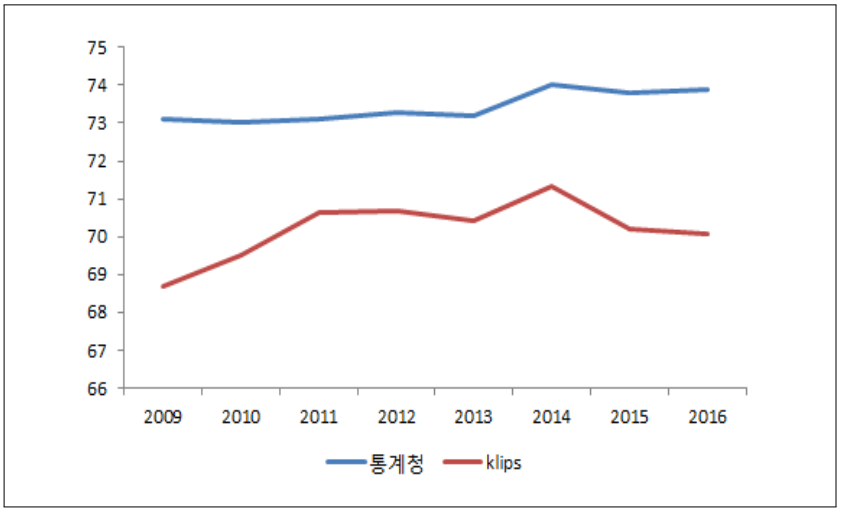
		2009	2010	2011	2012	2013	2014	2015	2016
취업자	KLIPS	13,388	13,393	14,282	14,439	14,553	14,920	14,870	14,881
	통계청	13,734	13,915	14,153	14,387	14,573	14,839	14,971	15,122
	KLIPS/통계청	0.97	0.96	1.01	1.00	1.00	1.01	0.99	0.98
실업자	KLIPS	607	372	487	490	455	411	340	411
	통계청	584	577	530	504	498	548	572	596
	KLIPS/통계청	1.04	0.64	0.92	0.97	0.91	0.75	0.59	0.69
비경활인구	KLIPS	6,375	6,035	6,136	6,188	6,304	6,163	6,454	6,525
	통계청	5,278	5,356	5,393	5,437	5,512	5,409	5,519	5,564
	KLIPS/통계청	1.21	1.13	1.14	1.14	1.14	1.14	1.17	1.17
경제활동참가율	KLIPS	68.7	69.52	70.65	70.7	70.42	71.33	70.21	70.09
	통계청	73.1	73.0	73.1	73.3	73.2	74.0	73.8	73.9
	KLIPS/통계청	0.94	0.95	0.97	0.96	0.96	0.96	0.95	0.95
고용률	KLIPS	65.7	67.6	68.3	68.4	68.3	69.4	68.6	68.2
	통계청	70.1	70.1	70.5	70.8	70.8	71.4	71.1	71.1
	KLIPS/통계청	0.94	0.96	0.97	0.97	0.96	0.97	0.97	0.96
실업률	KLIPS	4.3	2.7	3.3	3.3	3.0	2.7	2.2	2.7
	통계청	4.1	4.0	3.6	3.4	3.3	3.6	3.7	3.8
	KLIPS/통계청	1.06	0.68	0.92	0.97	0.92	0.74	0.60	0.71

비하여 줄어드는 것으로 나타나고 있다. 예를 들면 2016년 취업자 KLIPS/통계청은 전체 근로자를 대상으로 한 <표 5-1>에서는 0.96인 반면 남성만을 대상으로 한 <표 5-2>에서는 0.98로 나타나고 있다. 또한 실업자 KLIPS/통계청은 <표 5-1>의 0.67에서 <표 5-2>의 0.69로 증가하고 있다.

하지만 KLIPS의 비경제활동인구는 오히려 남성의 경우 더 과대 포착하고 있다. 2016년 비경활인구 KLIPS/통계청은 전체 근로자를 대상으로 한 <표 5-1>에서는 1.11인 반면 남성만을 대상으로 한 <표 5-2>에서는 1.17로 나타나고 있다. 따라서 KILPS의 취업자, 실업자, 비경제활동인구의 추정에 대한 전반적인 패턴은 남성근로자와 전체 근로자 사이 큰 차이를 보이지는 않지만 취업자와 실업자의 추정은 좀 더 정확한 반면, 비경제활동인구는 오히려 더 과대 추정하고 있다고 말할 수 있을 것이다.

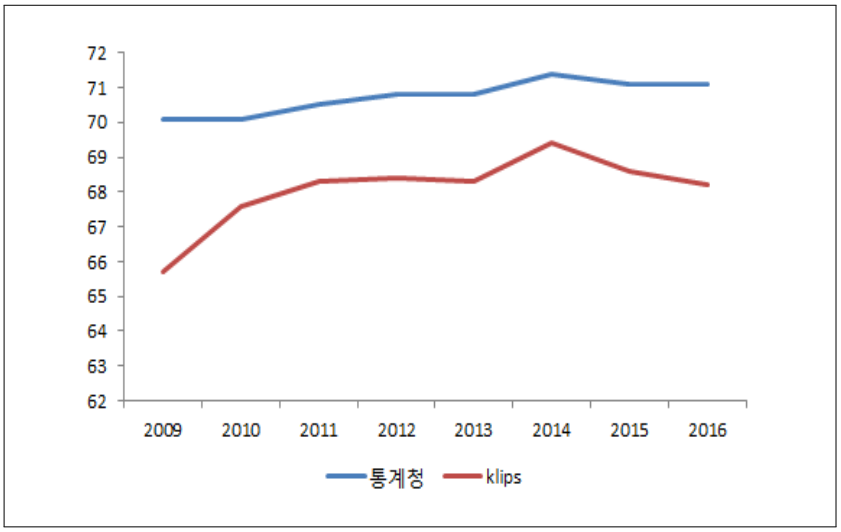
한편, 경제활동참가율, 고용률, 실업률을 보면 모두 KLIPS가 통계청의 경활조사보다 낮게 추정되고 있음을 알 수 있다. [그림 5-4]~[그림 5-6]은 이들 지표의 시계열을 보여주고 있는데 경제활동참가율과 고용률은 KLIPS와 경활조사 사이에 양의 상관관계를 가지는 한편, 실업률에 있어서는 상관관계가 상대적으로 낮게 나타나고 있다. 이는 전체 근로자를 대

〔그림 5-4〕 남성 경제활동참가율



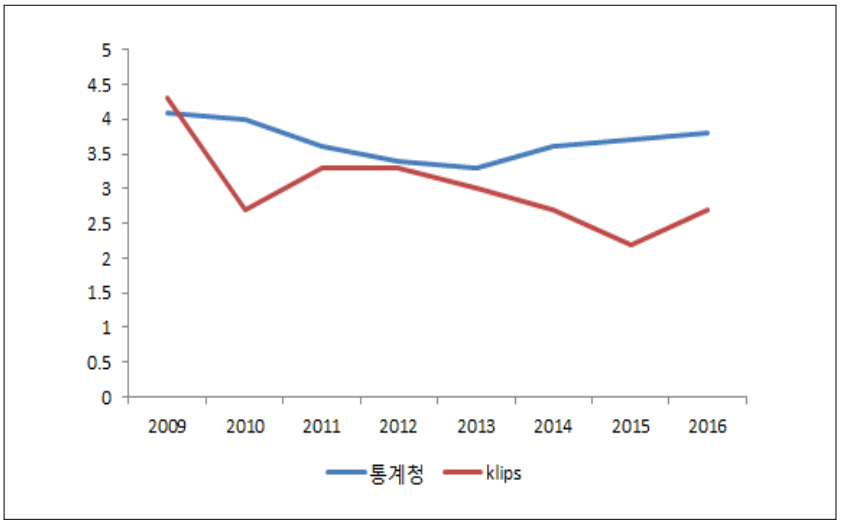
주 : 두 자료의 상관계수는 0.491로서 통계적으로 유의하지 않음.

〔그림 5-5〕 남성 고용률



주 : 두 자료의 상관계수는 0.813으로서 95% 수준에서 통계적으로 유의함.

〔그림 5-6〕 남성 실업률



주 : 두 자료의 상관계수는 0.258로서 90% 수준에서 통계적으로 유의하지 않음.

상으로 한 앞의 결과와 크게 다르지 않다. 다만, 경제활동참가율에 있어서 두 자료 사이의 상관관계가 상대적으로 낮게 나타나고 있음은 감안해

야 할 사항이다.

3. 여성 근로자

여성의 노동시장 활동에 대한 통계는 <표 5-3>에 제시된다. 2009년을 예외로 하면, 비경활인구를 제외한 모든 지표에서 KLIPS의 수치가 통계청의 경제활동인구조사에서 측정되는 수치보다 더 작게 나타나고 있음을 알 수 있다. 또한 매 지표의 세 번째 행에 있는 KLIPS/통계청의 수치를 보면 남성에 비하여 여성이 더 낮게 나타남으로써 KLIPS에서 여성이 남성보다 좀 더 과소 추정되고 있음을 알 수 있다. 하지만 비경활인구규모

〈표 5-3〉 OECD 기준 KLIPS, 통계청 여성인구(천 명) 총괄

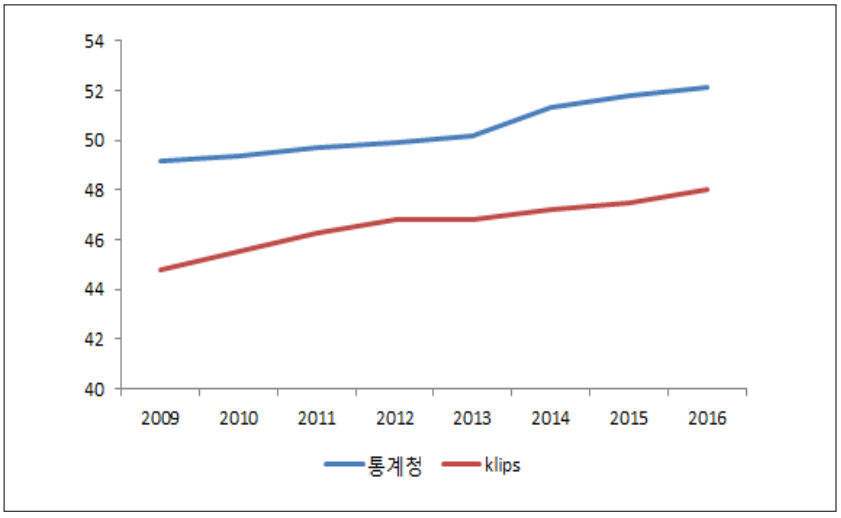
		2009	2010	2011	2012	2013	2014	2015	2016
취업자	KLIPS	8,776	9,038	9,517	9,659	9,868	9,974	10,063	10,336
	통계청	9,772	9,914	10,091	10,294	10,494	10,761	10,965	11,114
	KLIPS/통계청	0.90	0.91	0.94	0.94	0.94	0.93	0.92	0.93
실업자	KLIPS	444	250	251	316	205	283	343	268
	통계청	304	342	325	316	309	389	405	416
	KLIPS/통계청	1.46	0.73	0.77	1.00	0.66	0.73	0.85	0.64
비경활인구	KLIPS	11,360	11,116	11,335	11,353	11,464	11,475	11,508	11,482
	통계청	10,420	10,485	10,561	10,645	10,710	10,568	10,586	10,604
	KLIPS/통계청	1.09	1.06	1.07	1.07	1.07	1.09	1.09	1.08
경제활동참가율	KLIPS	44.8	45.52	46.3	46.8	46.8	47.2	47.5	48.0
	통계청	49.2	49.4	49.7	49.9	50.2	51.3	51.8	52.1
	KLIPS/통계청	0.91	0.92	0.93	0.94	0.93	0.92	0.92	0.92
고용률	KLIPS	42.6	44.3	45.1	45.3	45.8	45.9	45.9	46.8
	통계청	47.7	47.8	48.1	48.4	48.8	49.5	49.9	50.2
	KLIPS/통계청	0.89	0.93	0.94	0.94	0.94	0.93	0.92	0.93
실업률	KLIPS	4.8	2.7	2.6	3.2	2.0	2.8	3.3	2.5
	통계청	3.0	3.3	3.1	3.0	2.9	3.5	3.6	3.6
	KLIPS/통계청	1.61	0.82	0.83	1.06	0.70	0.79	0.92	0.70

의 경우에는 여성의 KLIPS/통계청의 수치가 남성보다 더 낮게 나타남으로써 이 지표에 있어서는 여성이 남성보다 더 정확하게 측정되고 있다고 말할 수 있다.

여성의 경제활동참가율, 고용률, 실업률의 시계열적 변화는 [그림 5-7]에서부터 [그림 5-9]에 제시되고 있는데, 남성의 경우와 비교하여 특징적으로 나타나는 점은 남성의 경우보다 여성의 경우에 KLIPS의 지표가 과소 추정되고 있는 반면, KLIPS의 시계열 추세의 경우에는 여성의 추세가 남성보다 훨씬 더 통계청의 경활조사와 유사하게 움직이고 있다는 것이다. 즉 절대적인 지표의 크기에 있어서는 남성이 여성보다 좀 더 정확하게 KLIPS에서 측정되지만, 지표의 추세에 있어서는 여성이 좀 더 정확하다고 볼 수 있다.

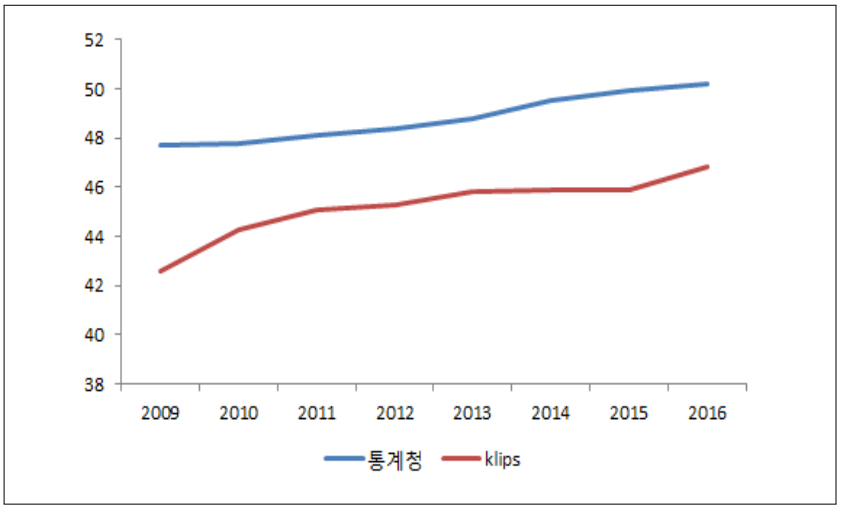
여성의 경제활동참가율의 경우 KLIPS와 통계청의 상관계수는 0.906으로 남성의 0.491보다 거의 두 배의 강한 상관관계를 보여주고 있으며 95% 수준에서도 통계적으로 유의하게 나타나고 있다. 여성의 고용률의 경우 KLIPS와 통계청의 상관계수는 0.842로 남성의 0.813보다 약간 높게 나타나고 있다. 따라서 KLIPS에서 여성의 고용률보다는 경제활동참가율

〔그림 5-7〕 여성 경제활동참가율



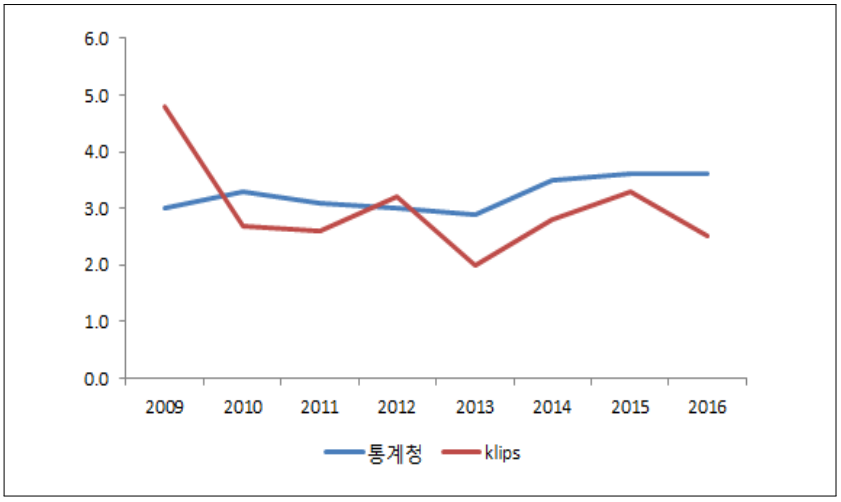
주 : 두 자료의 상관계수는 0.906으로서 95% 수준에서 통계적으로 유의함.

〔그림 5-8〕 여성 고용률



주 : 두 자료 간의 상관계수는 0.842로 95% 수준에서 통계적으로 유의함.

〔그림 5-9〕 여성 실업률



주 : 두 자료의 상관계수는 -0.133으로서 90% 수준에서 통계적으로 유의하지 않음.

의 추세를 좀 더 정확하게 반영하고 있다고 볼 수 있다.

한편, 여성의 실업률의 경우에는 KLIPS와 통계청의 상관계수는 -0.133으로 나타나고 있어 두 자료 사이에 음의 관계를 보이고 있다. 이는 적어

도 양의 관계를 보이는 남성의 경우보다 더 문제가 될 수 있다고 볼 수 있으나 두 경우 모두 90% 수준에서도 통계적으로 유의하지 않기 때문에 KLIPS에서 남성이 여성보다 더 정확하게 실업률의 변화를 반영한다고 말할 수는 없을 것이다.

제3절 월별 노동시장 활동 현황

통계청의 경제활동인구조사는 매월 조사가 이루어지기 때문에 월별 노동시장활동을 잘 반영하고 있다. 하지만 KLIPS는 원칙적으로 월별 자료를 구축하려는 것이 아니기 때문에 월별 노동시장활동을 알아보는 데는 무리가 있지만 설문의 인터뷰 시점(면접한 월)을 기준으로 월별 통계를 구축하는 것은 가능하다. 여기서는 경제활동참가율, 고용률, 실업률에 대하여 통계청의 경활조사와 KLIPS를 월별로 비교해 보기로 한다.

1. 전체 근로자

2009년부터 2016년까지 KLIPS의 월별 조사 빈도는 <표 5-4>와 [그림 5-10]과 같다. 2009~2016년 8년간 1월부터 12월까지 조사 횟수를 보면 7월부터 11월까지는 8년 동안 매달 조사가 이루어진 반면, 나머지 달은 해마다 조사가 이루어지는 것은 아님을 알 수 있다. 예를 들면, 1월은 8년 동안 두 해만 조사가 이루어졌고, 2월은 단 한 차례만 조사가 이루어졌다.

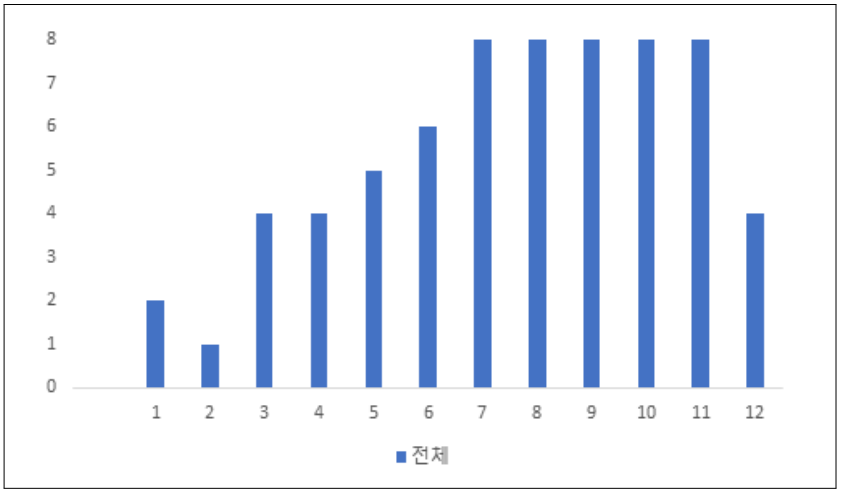
한편, 조사가 이루어진 달이라 하더라도 취업자, 실업자, 비경활인구에 대하여 골고루 이루어지는 것이 아니라 어떤 달에는 취업자만 조사되고, 어떤 달에는 실업자만 조사되는 경우도 있다. 예를 들어, 2009년 1월에는 조사가 이루어졌으나 취업자에 대해서만 조사가 이루어졌다.[38] 따라서 이런 달의 경우, KLIPS를 이용한 경제활동참가율, 취업률, 실업률에 대한 통계지표는 구축가능하지 않다.

38) <부표 5-4>를 참조할 것.

〈표 5-4〉 2009~2016년 월별 조사 항목 및 횟수-전체 근로자

월	취업자	실업자	비경제 활동 인구	경제활동 참가율	고용률	실업률	조사 횟수
1	2	0	1	0	0	0	2
2	1	1	1	1	1	1	1
3	4	1	2	1	1	1	4
4	4	4	4	4	4	4	4
5	4	4	5	4	4	4	5
6	6	6	6	6	6	6	6
7	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8
9	8	8	8	8	8	8	8
10	8	7	7	7	7	7	8
11	8	7	7	7	7	7	8
12	4	4	4	4	4	4	4

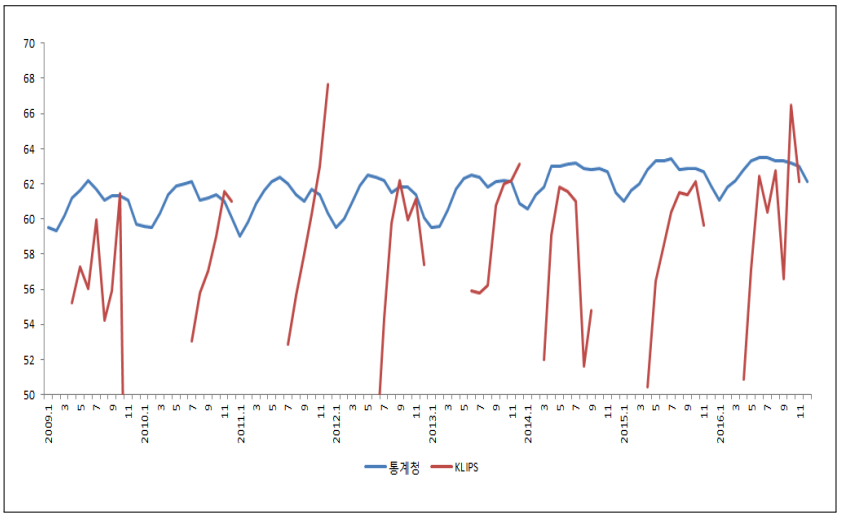
〔그림 5-10〕 2009~2016년 월별 조사 횟수-전체 근로자



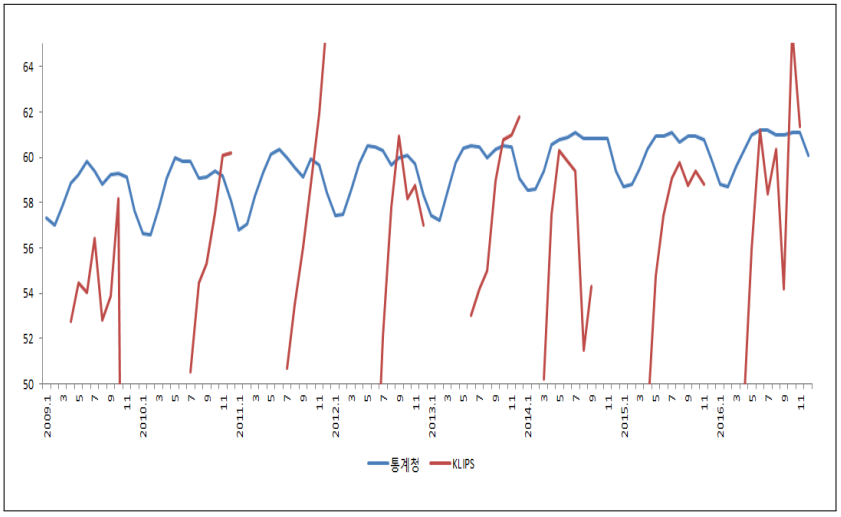
2009년 1월부터 2016년 12월까지 통계청의 경활조사와 한국노동연구원의 KLIPS의 월별 경제활동참가율, 취업률, 실업률은 <부표 5-4>에서부터 <부표 5-11>에 제시되고 있으며 이를 그래프로 나타낸 것이 [그림 5-11]

~[그림 5-13]이다. 통계청의 그래프는 연속적인 반면에 KLIPS의 그래프는 불연속적인 것은 앞서 언급한 바와 같이 KLIPS의 경우 매달 조사가 이루어지는 것이 아니기 때문이다. 그래프를 보면 알 수 있듯이 KLIPS의 조사는 상반기보다는 주로 하반기에 이루어지고 있음을 알 수 있다.

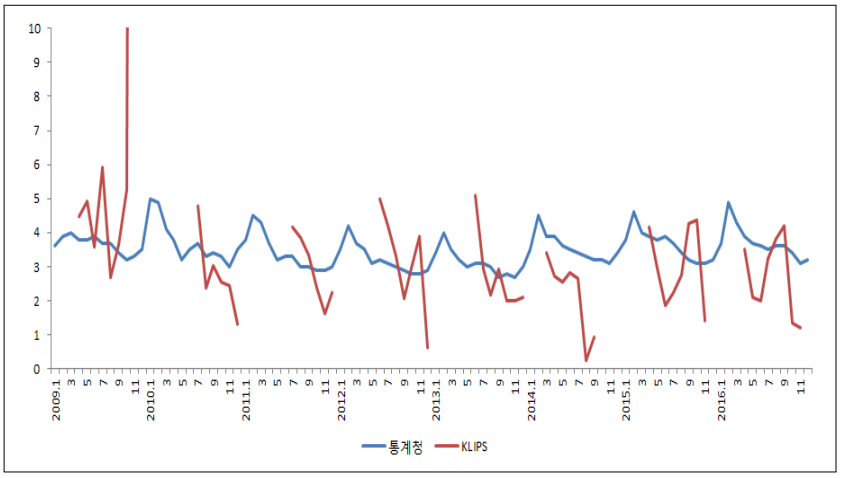
〔그림 5-11〕 2009~2016년 KLIPS와 통계청의 경제활동참가율-전체 근로자



〔그림 5-12〕 2009~2016년 KLIPS와 통계청의 고용률-전체 근로자



〔그림 5-13〕 2009~2016년 KLIPS와 통계청의 실업률-전체 근로자



우선 경제활동참가율을 보면 2011년 말과 2016년 말을 제외하고는 전반적으로 KLIPS가 통계청의 지표보다 낮은 수준에 있음을 알 수 있다. 이는 앞서 연도별 분석의 결과와 일치하는 것이다. 한편, KLIPS 월별 지표의 시계열이 불연속적이기 때문에 통계청 경제활동참가율의 증가추세를 KLIPS 월별 자료로 파악하기는 어렵다고 볼 수 있다. KLIPS에 나타난 월별 고용률 역시 경제활동참가율과 아주 비슷한 패턴을 보이고 있다. 즉 KLIPS의 월별 고용률은 전반적으로 통계청의 월별 고용률보다 낮은 수준을 보인다.

월별 실업률의 경우에는 KLIPS와 통계청의 경활조사 사이에 좀 더 불규칙한 패턴을 보이고 있다. 앞서 경제활동참가율, 고용률과의 차이점은 KLIPS의 실업률이 경활조사의 실업률을 초과하는 경우가 좀 더 많이 관찰된다는 것이다. 하지만 이러한 초과현상은 2013년 이전에 주로 나타나고, 그 이후에는 오히려 KLIPS 실업률이 경활조사 실업률보다 낮게 측정되는 경우가 더 빈번하게 나타나고 있다.

<표 5-5>는 월별 경제활동참가율, 고용률, 실업률에 대하여 KLIPS에서 구해진 값을 통계청 경활조사에서 얻어진 값으로 나눈 것이다. 만약 이 값이 1이면 두 자료에서 얻어진 지표가 동일한 것이며, 1보다 작으면 KLIPS가 과소 추정하는 것이며, 1보다 크면 과대 추정하는 것이다.

KLIPS의 월 평균은 경제활동참가율, 고용률, 실업률이 모두 관측된 달에 한하여 평균을 낸 것이다.[39)]

<표 5-5>는 KLIPS에서 월별 경제활동참가율, 고용률, 실업률이 모두 관찰되는 경우에는 KLIPS의 지표와 통계청 경활조사의 지표가 서로 크게 다르지 않음을 보여준다. 경제활동참가율의 경우에는 KLIPS/통계청의 값이 0.85에서 1.1 사이에 있는데 5월달 이후에는 그 값이 0.91에서 1.02 사이에 있어 비교적 오차가 작다고 할 수 있다.

고용률의 경우도 경제활동참가율과 유사한 패턴을 보이지만, 한 가지 차이점은 11월의 경우 KLIPS/통계청의 값이 0.88로 비교적 낮게 나타난다는 것이다. 실업률의 경우에도 11월은 상당한 문제를 보여주고 있다. KLIPS/통계청의 값이 3.64로 KLIPS의 실업률이 통계청 실업률의 거의 4배로 나타나는 것이다. 그 원인은 <부표 5-4>의 2009년 11월에 실업률이 너

〈표 5-5〉 2009~2016년 KLIPS/통계청 비율(월평균) - 전체 근로자

월	경제활동참가율	고용률	실업률
1	-	-	-
2	1.10	1.10	1.15
3	0.85	0.86	0.86
4	0.87	0.87	1.00
5	0.93	0.93	0.91
6	0.91	0.91	0.97
7	0.91	0.91	1.09
8	0.92	0.93	0.81
9	0.94	0.94	0.96
10	0.99	0.99	0.97
11	0.92	0.88	3.64
12	1.02	1.04	0.48

주 : 면접 월을 알 수 없는 경우는 missing으로 처리되고, 이들은 월평균 통계에서 제외되었음.

39) 2009년 1월의 경우 취업자에 대해서만 조사가 이루어졌기 때문에 경제활동참가율이나 실업률을 구할 수 없다. 따라서 2009년 1월은 1월 평균을 구하는 데 제외되었다.

무 과다하게 측정되었기 때문이다. 하지만 8월의 경우에도 KLIPS/통계청의 값이 0.81로 나타나기 때문에 경제활동참가율이나 취업률보다는 KLIPS의 실업률이 통계청 실업률과 더 큰 오차를 보인다고 말할 수 있을 것이다.

2. 남성 근로자

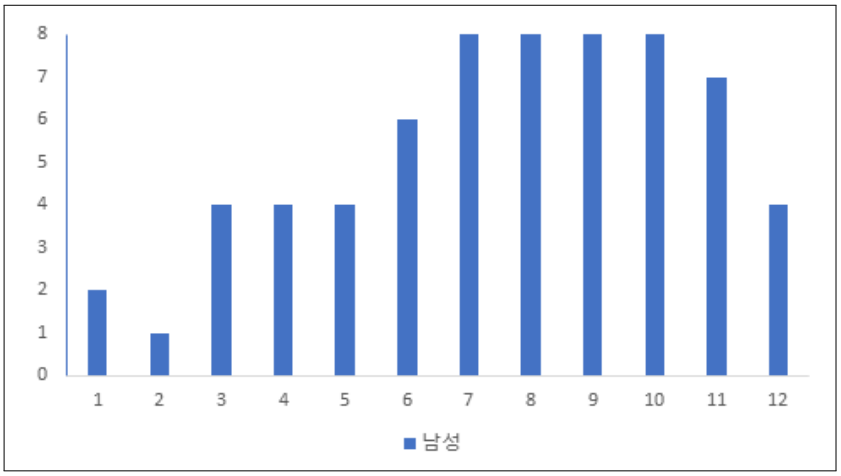
2009년부터 2016년까지 남성을 대상으로 한 KLIPS의 월별 조사 빈도는 <표 5-6>과 [그림 5-14]에 제시된다. 앞서 전체 근로자를 대상으로 <표 5-4>와의 차이점은 남성의 경우 5월과 11월에 조사 횟수가 한 번씩 적다는 것이다. 즉 이달에는 조사를 나가서 여성만을 대상으로 조사가 이루어졌다고 할 수 있다.

남성에 대한 조사 빈도는 상반기보다는 하반기에 더 높지만 12월은 예외적으로 4회에 그치고 있으며 특히 실업자에 대한 조사는 2회에 머물고 있다. 따라서 경제활동참가율, 고용률, 실업률에 대한 통계도 12월에는 2번만이 가능하다. 조사빈도가 낮은 상반기 경우에도 실업자에 대한 조사

〈표 5-6〉 2009~2016년 월별 조사 항목 및 횟수-남성

월	취업자	실업자	비경제 활동 인구	경제활동 참가율	고용률	실업률	조사 횟수
1	2	0	1	0	0	0	2
2	1	1	1	1	1	1	1
3	4	1	2	1	1	1	4
4	4	4	4	4	4	4	4
5	4	4	4	4	4	4	4
6	6	6	6	6	6	6	6
7	8	8	8	8	8	8	8
8	8	8	8	8	8	8	8
9	8	8	8	8	8	8	8
10	8	7	7	7	7	7	8
11	7	6	7	6	6	6	7
12	4	2	4	2	2	2	4

〔그림 5-14〕 2009~2016년 월별 조사 횟수 - 남성



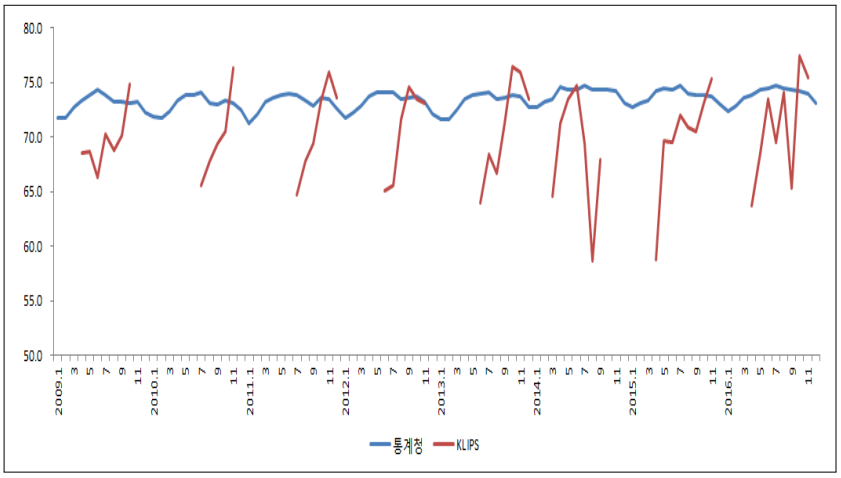
는 여전히 취업자에 비하여 작게 나타나고 있다.

남성을 대상으로 하여 2009년 1월부터 2016년 12월까지 통계청의 경활조사와 한국노동연구원의 KLIPS를 통하여 계산한 월별 경제활동참가율, 취업률, 실업률은 <부표 5-12>에서부터 <부표 5-19>에 제시되고 있으며 이를 그래프로 나타낸 것이 [그림 5-15]~[그림 5-17]이다.

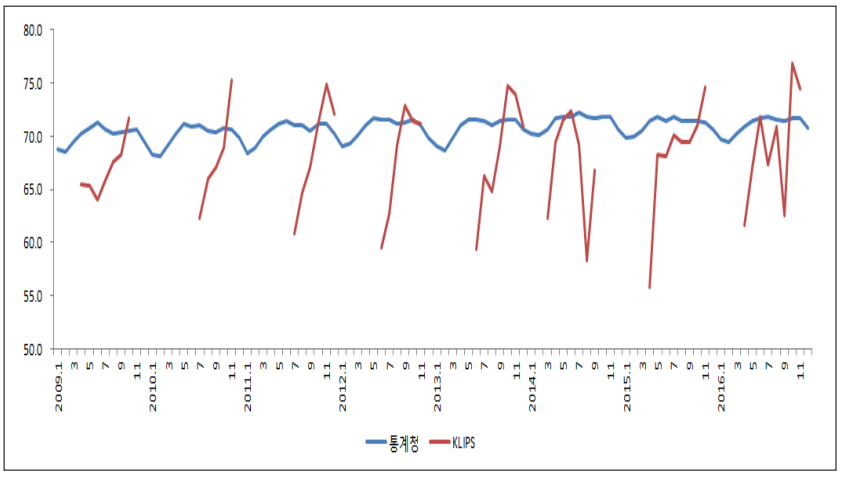
KILIPS의 월별 경제활동참가율을 보면 간헐적으로 통계청 경활조사의 경제활동참가율을 초과하는 달이 있지만 전반적으로는 KLIPS가 통계청의 지표보다 낮은 수준에 머물고 있음을 알 수 있다. 또한, 한 해의 상반기에는 통계청과 KLIPS의 격차가 크게 나타나지만 하반기로 갈수록 두 조사의 차이는 줄어드는 것을 보이고 있다. 이는 KLIPS의 조사가 하반기에 좀 더 많이 이루어지기 때문이다.

KLIPS에 나타난 월별 고용률도 경제활동참가율과 아주 비슷한 패턴을 보이고 있다. 즉 KLIPS의 월별 고용률은 전반적으로 통계청의 월별 고용률보다 낮은 수준을 보인다. 하지만 실업률의 경우에는 통계청과 KLIPS의 차이가 현저하게 나타남을 볼 수 있는데 조사가 많이 이루어지고 있는 하반기의 경우에도 이러한 차이는 크게 줄어드는 경향을 보이지 않고 있다. 따라서 KLIPS에서 구해지는 월별 남성 실업률과 통계청의 실업률은 상당한 괴리가 있다고 할 수 있다.

〔그림 5-15〕 2009~2016년 남성 KLIPS, 통계청 경제활동참가율

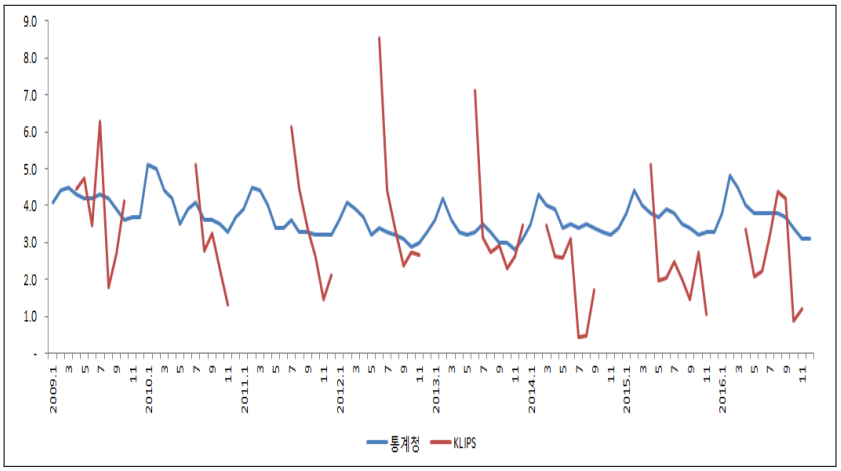


〔그림 5-16〕 2009~2016년 남성 KLIPS, 통계청 고용률



<표 5-7>은 앞의 <표 5-5>와 같은 요령으로 남성의 월별 경제활동참가율, 고용률, 실업률에 대하여 KLIPS에서 구해진 값을 통계청 경활조사에서 얻어진 값으로 나눈 것이다. KLIPS의 월평균은 경제활동참가율, 고용률, 실업률이 모두 관측된 달에 한하여 평균을 낸 것이다. 경제활동참가율은 KLIPS/통계청 비율이 0.88에서 1.02 사이에 있고, 고용률은 0.89에서 1.04 사이에, 실업률은 0.54에서 1.38 사이로 나타나고 있다. 특히

〔그림 5-17〕 2009~2016년 남성 KLIPS, 통계청 실업률



〈표 5-7〉 2009~2016년 KLIPS/통계청 비율(월평균) - 남성 근로자

월	경제활동참가율	고용률	실업률
1	-	-	-
2	0.99	0.97	1.38
3	0.88	0.89	0.84
4	0.89	0.89	1.00
5	0.95	0.95	0.80
6	0.93	0.92	1.20
7	0.92	0.92	1.05
8	0.93	0.93	0.78
9	0.95	0.96	0.81
10	1.01	1.01	0.77
11	1.02	1.04	0.54
12	1.01	1.02	0.84

주 : 면접 월을 알 수 없는 경우는 missing으로 처리되고, 이들은 월평균 통계에서 제외되었음.

경제활동참가율과 고용률은 5월 이후에는 최소값이 0.92로 올라서 이들 달의 경우에는 KLIPS의 월별 지표가 어느 정도 정확한 통계수치를 제공한다고 말할 수 있을 것이다. 하지만 KLIPS 실업률의 경우에는 어떤 특

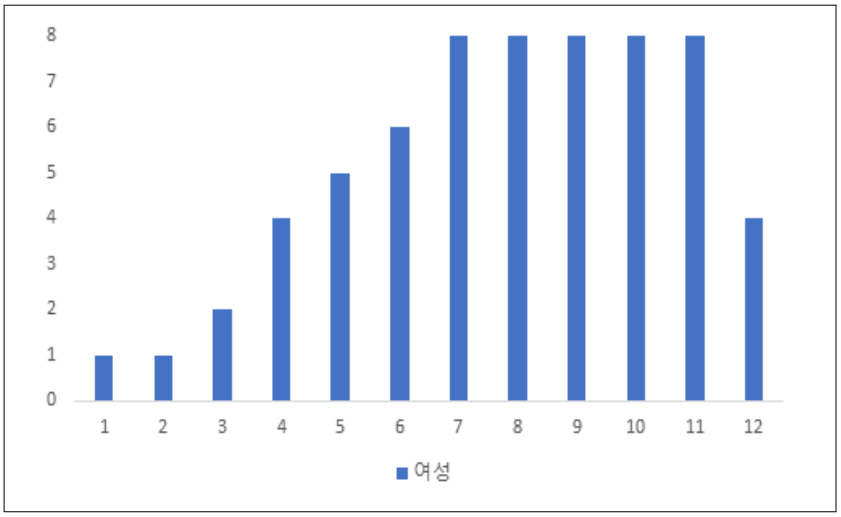
정한 패턴을 발견하기 어렵다. 상반기에는 0.8에서 1.38까지의 변화를 보이며, 하반기에도 0.54에서 1.05까지 나타나고 있다. 4월과 7월이 비교적 정확하게 실업률을 측정하고 있지만 이러한 경향이 앞으로도 지속되리라고 생각하기는 어렵다. 따라서 KLIPS를 이용하여 남성의 월별 실업률을 구하는 것은 주의가 요청된다.

3. 여성 근로자

여성을 대상으로 한 KLIPS의 월별 조사 빈도는 <표 5-8>과 [그림 5-18]에 지시되는데, 2009~2016년 8년 동안 1월과 2월은 단 1회만 조사된 것으로 나타나고 있다. 3월의 경우 남성은 4회가 조사된 반면 여성은 2회의 조사에 그치고 있다. 남성보다 조사 횟수가 많은 달은 11월 단 한 달이며 남성은 7회, 여성은 8회 모두 조사되고 있다. 따라서 남성에 비하여 여성의 조사 횟수의 월별 차이가 더 심하다고 말할 수 있다.

여성의 월별 경제활동참가율은 [그림 5-19]에 제시되고 있다. 남성과 유사하게 KLIPS의 경제활동참가율이 전반적으로 통계청 경활조사에 나타난 경제활동참가율보다 낮게 나타나고 있다. 하지만 남성과의 차이점

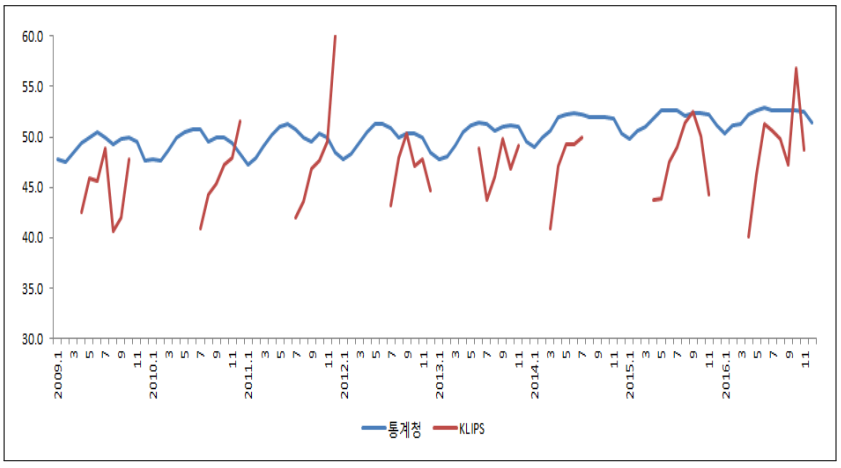
〔그림 5-18〕 2009~2016년 월별 조사 횟수 - 여성



〈표 5-8〉 2009~2016년 월별 조사 항목 및 횟수-여성

월	취업자	실업자	비경제 활동 인구	경제활동 참가율	고용률	실업률	조사 횟수
1	1	0	1	0	0	0	1
2	1	0	1	0	0	0	1
3	2	1	2	1	1	1	2
4	4	4	4	4	4	4	4
5	4	4	5	4	4	4	5
6	6	5	6	5	5	5	6
7	8	7	8	7	7	7	8
8	8	7	8	7	7	7	8
9	8	7	8	7	7	7	8
10	8	7	7	7	7	7	8
11	7	7	7	6	6	6	8
12	4	3	4	3	3	3	4

〔그림 5-19〕 2009~2016년 여성 KLIPS, 통계청 경제활동참가율

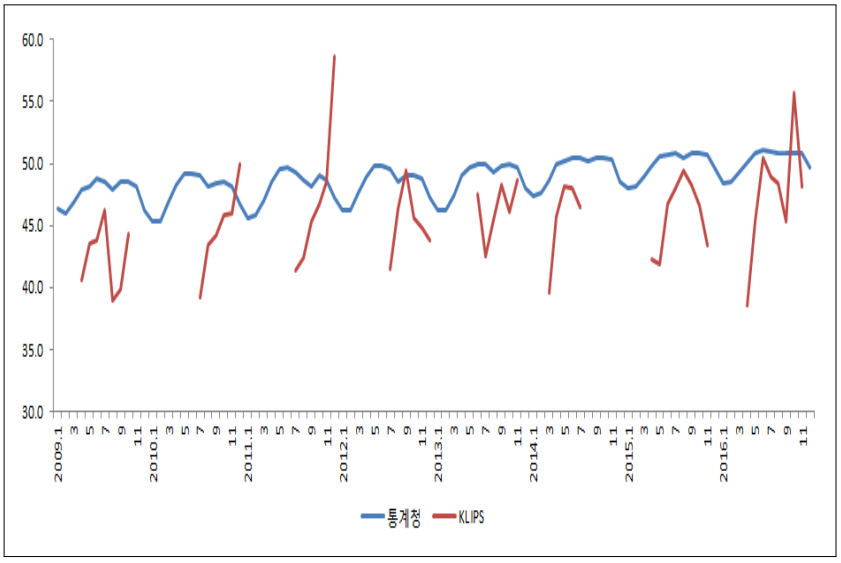


은 KLIPS의 경제활동참가율이 통계청의 경제활동참가율을 초과하는 달이 여성에게는 상대적으로 적게 나타나고 있다는 것이다. 따라서 연평균으로 보면 KLIPS에 나타난 여성의 경제활동참가율이 남성의 경제활동참

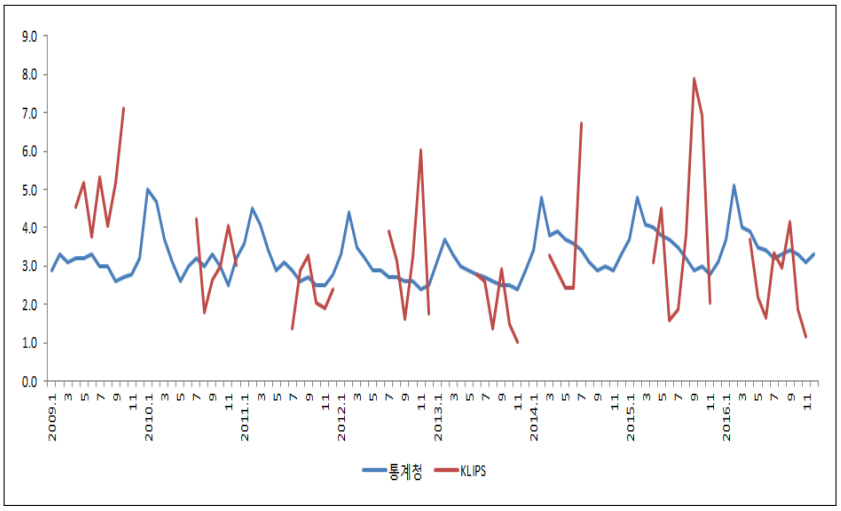
가율보다 더 과소하게 추정되는 경향을 보인다. 2009년부터 2016년까지 통계청의 경제활동참가율은 큰 폭은 아니지만 점진적으로 증가하고 있는 모습을 보이고 있다. KLIPS의 경우 그래프가 비연속적이기 때문에 추세를 확연히 볼 수는 없지만 전반적으로 그래프가 상향하고 있음을 알 수 있다.

고용률은 [그림 5-20]에 제시되고 있는데 그 패턴은 경제활동참가율과 크게 다르지 않다. KLIPS의 고용률은 통계청 경활조사의 고용률보다 전반적으로 낮게 나타나고 있으며, 통계청 고용률과 마찬가지로 KLIPS의 고용률도 2009년부터 점진적으로 증가하는 추세를 보여준다. 한편, [그림 5-21]은 월별 여성 실업률을 나타내는데 KLIPS에 나타난 여성 실업률은 통계청 여성 실업률과 비교하여 특정 달에 큰 차이를 보이고 있다. 경제활동참가율, 고용률과는 달리 실업률의 경우에는 KLIPS가 과대 추정하는 경향을 보이고 있는데 이러한 경향은 최근까지도 이어지고 있다. 남성과 마찬가지로 KLIPS로부터 구해진 여성 실업률은 그 크기에 있어서나 추세에 있어서 통계청 실업률과는 상당한 격차를 가진다고 할 수 있을 것이다.

〔그림 5-20〕 2009~2016년 여성 KLIPS, 통계청 고용률



〔그림 5-21〕 2009~2016년 여성 KLIPS, 통계청 실업률



마지막으로, <표 5-9>는 2009~2016년 여성의 월별 KLIPS 지표와 통계청 지표의 비율을 제시한다. 1월과 2월은 조사는 각각 1회씩 이루어졌지만 취업자, 실업자, 비경활인구 모두를 조사하지 않았기 때문에 경제활동참가율, 고용률, 실업률을 계산할 수 없다. 따라서 월별 평균은 3월부터나 가능하다.

먼저 경제활동참가율을 보면 KLIPS/통계청이 최소 0.82에서 최고 1.05로 남성보다는 변화폭이 다소 큰 것으로 나타났다.[40] 고용률의 경우에는 0.83부터 1.06으로 경제활동참가율과 크게 다르지 않다. 2009년부터 2016년까지 KLIPS에서 8회 모두 조사가 이루어진 달은 7월부터 11월인데 이 기간에는 0.89에서 0.94로 통계청 지표보다 모두 낮게 나타나고 있다. 따라서 KLIPS는 여성의 경제활동참가율과 고용률을 다소 과소 추정하고 있다고 말할 수 있을 것이다.

한편, 실업률의 경우에는 KLIPS의 모든 조사가 이루어진 7월부터 11월까지 최소 0.85에서 최대 1.30으로 변하고 있다. 따라서 여성 실업률의 경우에는 KLIPS가 항상 과소 추정하고 있다고 말할 수는 없으나 그 오차는 더 크다는 것을 알 수 있다.

40) 남성의 경우에는 해당 비율이 0.88부터 1.02이다.

〈표 5-9〉 2009~2016년 KLIPS/통계청 비율(월평균) - 여성 근로자

월	경제활동참가율	고용률	실업률
1	-	-	-
2	-	-	-
3	0.82	0.83	0.44
4	0.85	0.85	1.03
5	0.90	0.90	1.12
6	0.94	0.95	0.63
7	0.90	0.89	1.19
8	0.91	0.91	0.85
9	0.94	0.93	1.21
10	0.96	0.95	1.30
11	0.94	0.94	1.01
12	1.05	1.06	0.59

주 : 면접 월을 알 수 없는 경우는 missing으로 처리되고, 이들은 월평균 통계에서 제외되었음.

제4절 KLIPS와 통계청의 격차에 대한 원인 탐색

<표 5-1>에서 우리는 통계청의 경제활동참가율, 고용률, 실업률과 KLIPS의 해당 지표 사이에 차이가 있음을 발견하였다. 또한 월별 분석에서는 연도별 통계치보다 그 차이가 훨씬 더 큼을 발견하였다. 여기서는 그 이유에 대하여 탐색하여 보기로 한다.

우선 통계청의 고용통계를 보면 월별 통계와 해당 연도의 통계치 사이에 아주 큰 차이를 보이지는 않고 있다. 월에 따른 계절별 변동(seasonality)이 존재하기는 하지만 연평균과 아주 큰 차이를 보인다고 말하기는 어렵다. 따라서 만약 KLIPS의 설문응답자가 모두 동일한 달에 설문에 응답한다면 그 해의 통계청의 수치와 근접할 것으로 생각해 볼 수 있을 것이다. 이를 확인하기 위하여 2009~2016년의 연도별 응답 월에 대한 분산을 구해 보았다. 그 수치는 <표 5-10>의 각 셀 하단에 제시되었다.

〈표 5-10〉 KLIPS와 통계청 기준 경제활동지표(OECD 기준)

		2009	2010	2011	2012	2013	2014	2015	2016
경제활동 참가율	KLIPS	56.7	57.3	58.4	58.7	58.5	59.2	58.8	59.0
	통계청	60.8	61.0	61.1	61.3	61.5	62.4	62.6	62.8
	|KLIPS-통계청|	4.1	3.7	2.7	2.6	3.0	3.2	3.8	3.8
고용률	KLIPS	54.1	55.8	56.7	56.8	57.0	57.6	57.2	57.4
	통계청	58.6	58.7	59.1	59.4	59.5	60.2	60.3	60.4
	|KLIPS-통계청|	4.5	2.9	2.4	2.6	2.5	2.6	3.1	3.0
실업률	KLIPS	4.5	2.7	3.0	3.2	2.6	2.7	2.7	2.6
	통계청	3.6	3.7	3.4	3.2	3.1	3.5	3.6	3.7
	|KLIPS-통계청|	0.9	1.0	0.4	0.0	0.5	0.8	0.9	1.1
KLIPS 인터뷰 월의 표준편차		1.545	1.161	1.253	1.308	1.490	1.259	1.597	1.550

주 : KLIPS는 횡단면 가중치 사용. 통계청은 KOSIS 경제활동인구조사 연자료.

응답 월의 표준편차와 KLIPS와 통계청의 격차와의 상관관계를 계산한 결과, 경제활동참가율에 있어서는 0.508, 고용률은 0.500, 실업률은 0.306으로 정의 관계를 갖는 것으로 추정되었다. 즉 KLIPS의 응답 월의 분산이 커질수록 KLIPS와 통계청의 격차가 커진다고 할 수 있다. 이는 앞서의 예측, 즉 모든 사람이 동일한 달에 응답할수록 KLIPS의 값이 통계청의 값에 근접할 것이라는 것을 어느 정도 뒷받침한다고 할 수 있을 것이다. 다만, 상관계수의 값이 그리 높지 않아 95% 수준에서 통계적 유의성을 보이지는 않는다. 따라서 이 부분은 추후 KLIPS의 웨이브가 더 축적될 때 재검토될 필요가 있다.

KLIPS의 응답자가 얼마나 동일한 달에 응답하는지를 살펴보기 위하여 2개년 패널을 구축하여 동일 달 응답률과 다른 달 응답률을 계산해 보았다. 그 결과는 <표 5-11>에 제시되는데, 70% 이상의 응답자가 2개년 사이 다른 달에 응답하는 것으로 나타나고 있다. 이는 응답자의 연중 응답 시기가 매년 달라진다는 것을 의미한다.

만약 KLIPS의 응답자가 2개년 사이에 동일한 달에 응답할수록 KLIPS와 통계청의 격차가 줄어든다는 가설을 세울 수 있다. 이를 검증하기 위하여 동일한 달에 응답한 비율과 KLIPS와 통계청의 평균격차와의 상관

〈표 5-11〉 2개년 패널을 통한 동일 달, 다른 달 응답률

	표본수	동일 달 응답률 (%)	다른 달 응답률 (%)	경제활동참가율 평균격차	고용률 평균격차	실업률 평균격차
2009~2010	15,776	2.69	97.31	3.90	3.70	0.95
2010~2011	15,290	30.33	69.67	3.20	2.65	0.70
2011~2012	15,099	19.79	80.21	2.65	2.50	0.20
2012~2013	15,060	27.48	72.52	2.80	2.55	0.25
2013~2014	14,477	0.79	99.21	3.10	2.55	0.65
2014~2015	14,772	10.09	89.91	3.50	2.85	0.85
2015~2016	15,034	26.75	73.25	3.80	3.05	1.00

주 : 평균격차는 2개년 격차의 평균을 의미함.

계수를 구하여 보았다. 경제활동참가율에 있어서는 -0.291, 고용률은 -0.384, 실업률은 -0.298로 음의 관계를 갖는 것으로 추정되었다. 즉 KLIPS에서 2개년 동안 동일한 달에 응답할수록 KLIPS와 통계청의 격차가 감소한다고 말할 수 있다. 이를 통하여 동일한 달에 응답할수록 KLIPS의 값이 통계청의 값에 근접할 것이라고 예상할 수는 있을 것이다. 하지만, 앞에서와 같이 여기서도 상관계수의 값이 95% 수준에서 통계적 유의성을 보이지 않기 때문에 현재로선 가능성을 제시하는 것에 만족할 수밖에 없다.

제5절 요약 및 제언

본 연구는 한국노동연구원의 노동패널(KLIPS)에서 측정되는 노동시장 활동 지표와 통계청의 경제활동인구조사에서 나타나는 지표를 시계열적으로 비교함으로써 이들의 차이 정도를 확인하고 노동패널의 한계점과 이를 개선하기 위한 제언을 제시하고자 시도되었다. 노동시장 활동 지표로는 경제활동참가율, 고용률, 실업률(OECD 기준)을 사용하였으며 2009년부터 2016년까지 연도별, 월별로 이들 지표를 비교하였다. 본 연구의

결과는 다음과 같이 요약된다.

1) 연도별로 경제활동참가율, 고용률, 실업률(OECD 기준)을 비교한 결과, KLIPS의 지표가 전반적으로 통계청의 지표보다 낮게 추정되고 있다. 또한, 이러한 결과는 남성과 여성을 구분하여도 큰 차이를 보이지 않는다.

2) KLIPS의 경제활동참가율과 고용률이 통계청의 해당 지표보다는 낮게 추정되지만 이들 사이의 상관관계는 양의 값을 가지며, 특히 여성의 경우 매우 높은 통계적 유의성을 보이고 있다. 따라서 여성의 경우 KLIPS의 경제활동참가율과 고용률의 변화는 우리나라 여성의 경제활동참가율과 고용률을 유추하는 데 유용하게 사용될 수 있다.

3) KLIPS의 실업률은 통계청의 실업률과 상당한 괴리를 보이며 KLIPS 실업률과 통계청 실업률은 남성, 여성 모두 낮은 상관관계를 가지며 90% 수준에서도 통계적으로도 유의하지 않게 나타나고 있다.

4) KLIPS와 통계청의 격차는 응답자의 응답 월의 분산이 클수록, 2개년 패널의 응답 시기가 다를수록 크게 나타나고 있다.

5) KLIPS의 월별 조사 횟수를 보면 상반기보다는 하반기에 조사 횟수가 높으며, 따라서 하반기 KLIPS 지표가 통계청 지표와 좀 더 유사한 것으로 나타나고 있다.

6) KLIPS의 월별 경제활동참가율, 고용률은 2009년에는 통계청보다 주로 높게 추정되었으나 최근에 들어오면서는 오히려 낮게 추정되는 경향을 보이고 있다. 월별 실업률의 경우에는 KLIPS와 통계청 경활조사 사이에 어떤 특정한 패턴을 보이지 않고 있다.

이상의 결과를 토대로 KLIPS 조사의 개선방안을 생각해 보면 다음과 같다.

1) 비경제활동인구가 과도하게 조사되고 있는지 확인할 필요가 있다. 비경제활동인구가 과도하게 표집되면 경제활동참가율이 낮아지는 결과를 초래한다. 이러한 현상은 특히 남성에서 더 심하게 나타나고 있다.

2) 비경제활동인구로 응답한 사람과 실업자로 응답한 사람의 구분을 좀 더 정확히 할 필요가 있다. 이는 경제활동참가율과 실업률에 영향을 줄 수 있기 때문이다.

3) 취업자가 과소 표집되고 있는데, 특히 여성에게 이러한 현상이 더 많

이 나타나고 있다. 이러한 이유로 KLIPS의 고용률이 통계청의 고용률보다 낮게 나타난다.

4) 현재로선 KLIPS의 월별 노동시장활동 통계는 통계청과 차이를 보이기 때문에 월별지표 분석에 있어서는 주의를 요할 필요가 있다.

5) 패널조사의 특성상 응답자의 응답시기를 매년 같은 일자로 특정하는 것은 불가능하다. 따라서 이의 대안으로 한국노동패널의 조사시작 시기를 가능한 한 앞당기고, 응답시기의 분포를 최소화하여 응답기간이 특정 시기로 집중되도록 하는 방안을 고려해 볼 수 있다. 이렇게 할 경우 최소한 이들 시기에 대해서는 좀 더 정확한 노동시장 지표가 구축될 수 있을 것이다.

〈부표 5-1〉 KLIPS와 통계청 기준 경제활동지표(ILO 기준)

(단위 : 천 명, %)

		2009	2010	2011	2012	2013	2014	2015	2016
취업자	KLIPS	22,163	22,431	23,799	24,098	24,421	24,894	24,933	25,217
	통계청	23,506	23,829	24,244	24,681	25,066	25,599	25,936	26,235
	KLIPS/통계청	0.94	0.94	0.98	0.98	0.97	0.97	0.96	0.96
실업자	KLIPS	936	551	616	681	552	587	537	559
	통계청	829	832	759	723	720	834	-	-
	KLIPS/통계청	1.13	0.66	0.81	0.94	0.77	0.70	-	-
비경활 인구	KLIPS	17,851	17,222	17,594	17,667	17,877	17,746	18,108	18,127
	통계청	15,758	15,929	16,048	16,179	16,310	16,080	-	-
	KLIPS/통계청	1.13	1.08	1.10	1.09	1.10	1.10	-	-
경제활동 참가율	KLIPS	56.4	57.2	58.1	58.4	58.3	59.0	58.4	58.7
	통계청	60.7	60.8	60.9	61.1	61.3	62.2	-	-
	KLIPS/통계청	0.93	0.94	0.95	0.96	0.95	0.95	-	-
고용률	KLIPS	54.1	55.8	56.7	56.8	57	57.6	57.2	57.4
	통계청	58.6	58.7	59.1	59.4	59.5	60.2	60.3	60.4
	KLIPS/통계청	0.92	0.95	0.96	0.96	0.96	0.96	0.95	0.95
실업률	KLIPS	4.1	2.4	2.5	2.8	2.2	2.3	2.1	2.2
	통계청	3.4	3.4	3	2.8	2.8	3.2	-	-
	KLIPS/통계청	1.21	0.71	0.83	1.00	0.79	0.72	-	-

주 : KLIPS는 횡단면 가중치 사용. 통계청은 KOSIS 경제활동인구조사 연자료.

〈부표 5-2〉 남성 KLIPS와 통계청 기준 경제활동지표(ILO 기준)

(단위 : 천 명, %)

		2009	2010	2011	2012	2013	2014	2015	2016
취업자	KLIPS	13,388	13,393	14,282	14,439	14,553	14,920	14,870	14,881
	통계청	13,734	13,915	14,153	14,387	14,573	14,839	14,971	15,122
	KLIPS/통계청	0.97	0.96	1.01	1.00	1.00	1.01	0.99	0.98
실업자	KLIPS	548	337	429	427	371	338	282	336
	통계청	553	535	480	448	449	496	-	-
	KLIPS/통계청	0.99	0.63	0.89	0.95	0.83	0.68	-	-
비경활 인구	KLIPS	6,434	6,070	6,194	6,251	6,388	6,236	6,512	6,599
	통계청	5,309	5,398	5,443	5,493	5,561	5,461	-	-
	KLIPS/통계청	1.21	1.12	1.14	1.14	1.15	1.14	-	-
경제활동 참가율	KLIPS	68.4	69.3	70.4	70.4	70.0	71.0	69.9	69.8
	통계청	72.9	72.8	72.9	73.0	73.0	73.7	-	-
	KLIPS/통계청	0.94	0.95	0.97	0.96	0.96	0.96	-	-
고용률	KLIPS	65.7	67.6	68.3	68.4	68.3	69.4	68.6	68.2
	통계청	70.1	70.1	70.5	70.8	70.8	71.4	71.1	71.1
	KLIPS/통계청	0.94	0.96	0.97	0.97	0.96	0.97	0.97	0.96
실업률	KLIPS	3.9	2.5	2.9	2.9	2.5	2.2	1.9	2.2
	통계청	3.9	3.7	3.3	3.0	3.0	3.2	-	-
	KLIPS/통계청	1.01	0.66	0.88	0.96	0.83	0.69	-	-

주 : KLIPS는 횡단면 가중치 사용. 통계청은 KOSIS 경제활동인구조사 연자료.

〈부표 5-3〉 여성 KLIPS와 통계청 기준 경제활동지표(ILO 기준)

(단위 : 천 명, %)

		2009	2010	2011	2012	2013	2014	2015	2016
취업자	KLIPS	8,776	9,038	9,517	9,659	9,868	9,974	10,063	10,336
	통계청	9,772	9,914	10,091	10,294	10,494	10,761	10,965	11,114
	KLIPS/통계청	0.90	0.91	0.94	0.94	0.94	0.93	0.92	0.93
실업자	KLIPS	387	214	187	253	180	249	255	222
	통계청	276	297	280	274	270	338	-	-
	KLIPS/통계청	1.40	0.72	0.67	0.93	0.67	0.74	-	-
비경활 인구	KLIPS	11,417	11,152	11,400	11,416	11,489	11,509	11,596	11,527
	통계청	10,449	10,531	10,606	10,686	10,749	10,619	-	-
	KLIPS/통계청	1.09	1.06	1.07	1.07	1.07	1.08	-	-
경제활동 참가율	KLIPS	44.5	45.3	46.0	46.5	46.7	47.0	47.1	47.8
	통계청	49.0	49.2	49.4	49.7	50.0	51.1	-	-
	KLIPS/통계청	0.91	0.92	0.93	0.94	0.93	0.92	-	-
고용률	KLIPS	42.6	44.3	45.1	45.3	45.8	45.9	45.9	46.8
	통계청	47.7	47.8	48.1	48.4	48.8	49.5	49.9	50.2
	KLIPS/통계청	0.89	0.93	0.94	0.94	0.94	0.93	0.92	0.93
실업률	KLIPS	4.2	2.3	1.9	2.6	1.8	2.4	2.1	1.9
	통계청	2.7	2.9	2.7	2.6	2.5	3.0	-	-
	KLIPS/통계청	1.57	0.80	0.71	0.98	0.72	0.81	-	-

주 : KLIPS는 횡단면 가중치 사용. 통계청은 KOSIS 경제활동인구조사 연자료.

〈부표 5-4〉 2009년 월별 경제활동(KLIPS vs 통계청) - 전체

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	2,611	2,611						39,865	22,861	848	16,156	59.5	57.3	3.6
2	-							39,899	22,742	924	16,233	59.3	57.0	3.9
3	106,788	94,365		12,423				39,937	23,110	952	15,875	60.2	57.9	4.0
4	11,361,234	5,993,783	281,740	5,085,711	55.2	52.8	4.5	39,978	23,524	933	15,521	61.2	58.8	3.8
5	13,330,214	7,256,188	375,388	5,698,638	57.3	54.4	4.9	40,027	23,720	938	15,369	61.6	59.3	3.8
6	5,918,247	3,197,554	119,038	2,601,655	56.0	54.0	3.6	40,078	23,967	960	15,151	62.2	59.8	3.9
7	4,719,286	2,661,841	166,936	1,890,509	59.9	56.4	5.9	40,125	23,828	928	15,369	61.7	59.4	3.7
8	2,805,279	1,481,051	40,722	1,283,506	54.2	52.8	2.7	40,168	23,620	905	15,643	61.1	58.8	3.7
9	2,031,098	1,094,153	41,571	895,374	55.9	53.9	3.7	40,202	23,805	826	15,571	61.3	59.2	3.4
10	639,406	371,951	20,733	246,722	61.4	58.2	5.3	40,236	23,856	799	15,581	61.3	59.3	3.2
11	28,841	2,941	5,202	20,698	28.2	10.2	63.9	40,277	23,806	819	15,652	61.1	59.1	3.3
12	-							40,316	23,229	834	16,253	59.7	57.6	3.5
missing	6,950	6,950												
sum	40,949,954	22,163,388	1,051,330	17,735,236	56.7		4.5	40,092	23,506	889	15,698	60.8	58.6	3.6

주 : 실업자는 OECD 기준이며, missing은 인터뷰 월이 표시되지 않은 경우임.
15세 이상 인구, 취업자, 실업자, 비경활인구의 경우 KLIPS의 단위는 명이고 통계청의 단위는 천 명.
KLIPS는 가중치 사용, 통계청은 KOSIS 경제활동인구조사 월별 자료.

〈부표 5-5〉 2010년 월별 경제활동(KLIPS vs 통계청) - 전체

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	-							40,386	22,865	1,216	16,305	59.6	56.6	5.0
2	-							40,420	22,867	1,169	16,384	59.5	56.6	4.9
3	-							40,455	23,377	1,005	16,073	60.3	57.8	4.1
4	-							40,490	23,924	934	15,632	61.4	59.1	3.8
5	3,248			3,248				40,533	24,306	793	15,434	61.9	60.0	3.2
6	-							40,573	24,280	878	15,415	62.0	59.8	3.5
7	2,777,795	1,403,401	70,283	1,304,111	53.1	50.5	4.8	40,615	24,301	931	15,383	62.1	59.8	3.7
8	3,809,916	7,521,603	183,049	6,105,264	55.8	54.5	2.4	40,651	24,005	831	15,815	61.1	59.1	3.3
9	7,783,951	4,304,942	133,844	3,345,165	57.0	55.3	3.0	40,681	24,054	857	15,770	61.2	59.1	3.4
10	12,125,441	6,974,638	182,651	4,968,152	59.0	57.5	2.6	40,715	24,172	832	15,711	61.4	59.4	3.3
11	3,242,410	1,948,349	48,529	1,245,532	61.6	60.1	2.4	40,755	24,109	737	15,909	61.0	59.2	3.0
12	459,316	276,470	3,706	179,140	61.0	60.2	1.3	40,802	23,684	853	16,265	60.1	58.0	3.5
missing	1,697	1,697												
sum	40,203,774	22,431,100	622,062	17,150,612	57.3	55.8	2.7	40,590	23,829	920	15,841	61.0	58.7	3.7

주 : 실업자는 OECD 기준이며, missing은 인터뷰 월이 표시되지 않은 경우임.
15세 이상 인구, 취업자, 실업자, 비경활인구의 경우 KLIPS의 단위는 명이고 통계청의 단위는 천 명.
KLIPS는 가중치 사용, 통계청은 KOSIS 경제활동인구조사 월별 자료.

〈부표 5-6〉 2011년 월별 경제활동(KLIPS vs 통계청) - 전체

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	40,118	16,584		23,534				40,843	23,196	918	16,729	59.0	56.8	3.8
2	71,556	45,265	2,409	23,882	66.6	63.3	5.1	40,879	23,336	1,095	16,448	59.8	57.1	4.5
3	-							40,918	23,846	1,073	15,999	60.9	58.3	4.3
4	-							40,952	24,303	936	15,713	61.6	59.3	3.7
5	-							41,003	24,661	819	15,523	62.1	60.1	3.2
6	-							41,032	24,752	839	15,441	62.4	60.3	3.3
7	822,313	416,856	18,105	387,352	52.9	50.7	4.2	41,087	24,636	837	15,614	62.0	60.0	3.3
8	16,089,123	8,608,614	344,034	7,136,475	55.6	53.5	3.8	41,119	24,495	762	15,862	61.4	59.6	3.0
9	8,643,939	4,845,796	167,460	3,630,683	58.0	56.1	3.3	41,141	24,318	758	16,065	61.0	59.1	3.0
10	8,861,734	5,210,576	126,991	3,524,167	60.2	58.8	2.4	41,169	24,673	736	15,760	61.7	59.9	2.9
11	6,976,192	4,322,205	71,437	2,582,550	63.0	62.0	1.6	41,208	24,589	730	15,889	61.4	59.7	2.9
12	503,537	333,083	7,628	162,826	67.7	66.1	2.2	41,273	24,125	754	16,394	60.3	58.5	3.0
missing	-													
sum	42,008,512	23,798,979	738,064	17,471,469	58.4	56.7	3.0	41,052	24,244	855	15,953	61.1	59.1	3.4

주 : 실업자는 OECD 기준이며, missing은 인터뷰 월이 표시되지 않은 경우임.
15세 이상 인구, 취업자, 실업자, 비경활인구의 경우 KLIPS의 단위는 명이고 통계청의 단위는 천 명.
KLIPS는 가중치 사용, 통계청은 KOSIS 경제활동인구조사 월별 자료.

〈부표 5-7〉 2012년 월별 경제활동(KLIPS vs 통계청) - 전체

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	-							41,321	23,732	853	16,736	59.5	57.4	3.5
2	-							41,369	23,783	1,042	16,544	60.0	57.5	4.2
3	-							41,415	24,265	945	16,205	60.9	58.6	3.7
4	-							41,458	24,758	895	15,805	61.9	59.7	3.5
5	-							41,519	25,133	807	15,579	62.5	60.5	3.1
6	258,422	114,272	5,998	138,152	46.5	44.2	5.0	41,561	25,117	822	15,622	62.4	60.4	3.2
7	12,318,659	6,412,858	281,628	5,624,173	54.3	52.1	4.2	41,628	25,106	795	15,727	62.2	60.3	3.1
8	12,788,710	7,388,874	252,109	5,147,727	59.7	57.8	3.3	41,665	24,859	764	16,042	61.5	59.7	3.0
9	8,640,620	5,264,780	111,220	3,264,620	62.2	60.9	2.1	41,697	25,003	752	15,942	61.8	60.0	2.9
10	4,832,188	2812317	85319	1934552	60.0	58.2	2.9	41,730	25,069	718	15,943	61.8	60.1	2.8
11	2,789,679	1,639,026	66,539	1,084,114	61.1	58.8	3.9	41,774	24,941	710	16,123	61.4	59.7	2.8
12	815,905	465,300	2,951	347,654	57.4	57.0	0.6	41,848	24,402	737	16,709	60.1	58.3	2.9
missing	1,179	580		599										
sum	42,445,362	24,098,007	805,764	17,541,591	59.2	57.4	3.2	41,582	24,681	820	16,081	61.3	59.4	3.2

주 : 실업자는 OECD 기준이며, missing은 인터뷰 월이 표시되지 않은 경우임.
15세 이상 인구, 취업자, 실업자, 비경활인구의 경우 KLIPS의 단위는 명이고 통계청의 단위는 천 명.
KLIPS는 가중치 사용, 통계청은 KOSIS 경제활동인구조사 월별 자료.

〈부표 5-8〉 2013년 월별 경제활동(KLIPS vs 통계청) - 전체

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	-							41,876	24,054	847	16,975	59.5	57.4	3.4
2	-							41,927	23,984	990	16,953	59.6	57.2	4.0
3	-							41,961	24,514	883	16,564	60.5	58.4	3.5
4	-							42,002	25,103	825	16,074	61.7	59.8	3.2
5	-							42,048	25,398	797	15,853	62.3	60.4	3.0
6	2,621,385	1,390,657	74,397	1,156,331	55.9	53.1	5.1	42,098	25,478	813	15,807	62.5	60.5	3.1
7	11,633,493	6,303,886	189,089	5,140,518	55.8	54.2	2.9	42,145	25,473	828	15,844	62.4	60.4	3.1
8	9,084,313	4,997,226	111,125	3,975,962	56.2	55.0	2.2	42,168	25,291	783	16,094	61.8	60.0	3.0
9	8,099,851	4,776,459	143,583	3,179,809	60.7	59.0	2.9	42,194	25,466	720	16,008	62.1	60.4	2.7
10	7,502,167	4,560,209	92,683	2,849,275	62.0	60.8	2.0	42,218	25,545	724	15,949	62.2	60.5	2.8
11	2,735,777	1,668,102	33,725	1,033,950	62.2	61.0	2.0	42,246	25,530	700	16,016	62.1	60.4	2.7
12	1,172,550	724,246	15,513	432,791	63.1	61.8	2.1	42,273	24,962	774	16,537	60.9	59.0	3.0
missing														
sum	42,849,536	24,420,785	660,115	17,768,636	58.5	57.0	2.6	42,096	25,067	807	16,223	61.5	59.5	3.1

주 : 실업자는 OECD 기준이며, missing은 인터뷰 월이 표시되지 않은 경우임.
15세 이상 인구, 취업자, 실업자, 비경활인구의 경우 KLIPS의 단위는 명이고 통계청의 단위는 천 명.
KLIPS는 가중치 사용, 통계청은 KOSIS 경제활동인구조사 월별 자료.

〈부표 5-9〉 2014년 월별 경제활동(KLIPS vs 통계청) - 전체

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	-							42,304	24,759	891	16,654	60.6	58.5	3.5
2	-							42,343	24,819	1,178	16,346	61.4	58.6	4.5
3	6,489,081	3,258,840	114,582	3,115,659	52.0	50.2	3.4	42,378	25,163	1,024	16,191	61.8	59.4	3.9
4	13,264,230	7,625,539	213,004	5,425,687	59.1	57.5	2.7	42,414	25,684	1,030	15,700	63.0	60.6	3.9
5	12,161,602	7,329,983	190,818	4,640,801	61.8	60.3	2.5	42,453	25,811	951	15,691	63.0	60.8	3.6
6	8,268,604	4,947,975	143,495	3,177,134	61.6	59.8	2.8	42,490	25,875	949	15,666	63.1	60.9	3.5
7	1,717,789	1,020,662	27,815	669,312	61.0	59.4	2.7	42,528	25,979	912	15,637	63.2	61.1	3.4
8	707,667	364,377	944	342,346	51.6	51.5	0.3	42,572	25,885	890	15,797	62.9	60.8	3.3
9	590,480	320,591	3,021	266,868	54.8	54.3	0.9	42,611	25,917	849	15,845	62.8	60.8	3.2
10	19,290	19,290						42,646	25,951	858	15,837	62.9	60.9	3.2
11	4,565	4,565						42,688	25,968	818	15,902	62.7	60.8	3.1
12	-							42,728	25,384	886	16,458	61.5	59.4	3.4
missing	1,648	1,648												
sum	43,224,956	24,893,470	693,679	17,637,807	59.2	57.6	2.7	42,513	25,600	936	15,977	62.4	60.2	3.5

주 : 실업자는 OECD 기준이며, missing은 인터뷰 월이 표시되지 않은 경우임.
15세 이상 인구, 취업자, 실업자, 비경활인구의 경우 KLIPS의 단위는 명이고 통계청의 단위는 천 명.
KLIPS는 가중치 사용, 통계청은 KOSIS 경제활동인구조사 월별 자료.

〈부표 5-10〉 2015년 월별 경제활동(KLIPS vs 통계청) - 전체

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	-							42,775	25,106	988	16,681	61.0	58.7	3.8
2	-							42,827	25,195	1,203	16,429	61.6	58.8	4.6
3	501	501						42,875	25,501	1,076	16,298	62.0	59.5	4.0
4	2,227,776	1,076,956	46,712	1,104,108	50.4	48.3	4.2	42,924	25,900	1,053	15,971	62.8	60.3	3.9
5	11,262,327	6,171,518	189,229	4,901,580	56.5	54.8	3.0	42,975	26,189	1,022	15,764	63.3	60.9	3.8
6	8,149,757	4,679,630	88,813	3,381,314	58.5	57.4	1.9	43,025	26,205	1,050	15,770	63.3	60.9	3.9
7	10,568,915	6,241,528	143,544	4,183,843	60.4	59.1	2.2	43,054	26,305	998	15,751	63.4	61.1	3.7
8	7,034,023	4,205,498	119,171	2,709,354	61.5	59.8	2.8	43,086	26,141	923	16,022	62.8	60.7	3.4
9	1,676,906	985,330	43,807	647,769	61.4	58.8	4.3	43,120	26,264	866	15,990	62.9	60.9	3.2
10	1,565,297	929,632	42,459	593,206	62.1	59.4	4.4	43,149	26,298	839	16,012	62.9	60.9	3.1
11	1,091,932	642,125	9,305	440,502	59.7	58.8	1.4	43,181	26,253	829	16,099	62.7	60.8	3.1
12	-							43,216	25,879	868	16,469	61.9	59.9	3.2
missing	-													
sum	43,577,434	24,932,718	683,040	17,961,676	58.8	57.2	2.7	43,017	25,936	976	16,105	62.6	60.3	3.6

주 : 실업자는 OECD 기준이며, missing은 인터뷰 월이 표시되지 않은 경우임.
15세 이상 인구, 취업자, 실업자, 비경활인구의 경우 KLIPS의 단위는 명이고 통계청의 단위는 천 명.
KLIPS는 가중치 사용, 통계청은 KOSIS 경제활동인구조사 월별 자료.

〈부표 5-11〉 2016년 월별 경제활동(KLIPS vs 통계청) - 전체

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제활동인구	경제활동 참가율	고용률	실업률
1	-							43,239	25,445	988	16,806	61.1	58.8	3.7
2	-							43,274	25,418	1,317	16,540	61.8	58.7	4.9
3	3,445	3,445						43,313	25,800	1,155	16,358	62.2	59.6	4.3
4	5,456,477	2,680,177	97,530	2,678,770	50.9	49.1	3.5	43,348	26,153	1,075	16,120	62.8	60.3	3.9
5	13,638,794	7,629,132	164,914	5,844,748	57.1	55.9	2.1	43,387	26,450	1,005	15,932	63.3	61.0	3.7
6	9,784,987	5,988,945	121,938	3,674,104	62.5	61.2	2.0	43,421	26,559	1,004	15,859	63.5	61.2	3.6
7	8,117,166	4,741,358	159,255	3,216,553	60.4	58.4	3.2	43,442	26,603	975	15,864	63.5	61.2	3.5
8	3,871,751	2,336,801	93,237	1,441,713	62.8	60.4	3.8	43,465	26,528	996	15,941	63.3	61.0	3.6
9	1,083,400	587,260	25,727	470,413	56.6	54.2	4.2	43,488	26,531	986	15,972	63.3	61.0	3.6
10	1,319,019	865,164	11,673	442,182	66.5	65.6	1.3	43,512	26,577	923	16,013	63.2	61.1	3.4
11	627,235	384,872	4,651	237,712	62.1	61.4	1.2	43,537	26,592	854	16,091	63.0	61.1	3.1
12	-							43,561	26,168	867	16,526	62.1	60.1	3.2
missing	-													
sum	43,902,272	25,217,153	678,924	18,006,195	59.0	57.4	2.6	43,416	26,235	1,012	16,169	62.8	60.4	3.7

주 : 실업자는 OECD 기준이며, missing은 인터뷰 월이 표시되지 않은 경우임.
15세 이상 인구, 취업자, 실업자, 비경활인구의 경우 KLIPS의 단위는 명이고 통계청의 단위는 천 명.
KLIPS는 가중치 사용, 통계청은 KOSIS 경제활동인구조사 월별 자료.

〈부표 5-12〉 2009년 월별 경제활동(KLIPS vs 통계청) - 남성

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	2,611	2,611						19,484	13,405	573	5,516	71.7	68.8	4.1
2	-							19,501	13,358	615	5,526	71.7	68.5	4.4
3	67,934	58,504		9,430				19,518	13,545	638	5,335	72.7	69.4	4.5
4	5,536,892	3,627,147	169,045	1,740,700	68.6	65.5	4.5	19,538	13,735	617	5,188	73.4	70.3	4.3
5	6,650,382	4,347,202	216,062	2,087,118	68.6	65.4	4.7	19,564	13,851	607	5,109	73.9	70.8	4.2
6	2,984,826	1,911,212	68,662	1,004,952	66.3	64.0	3.5	19,589	13,967	612	5,011	74.4	71.3	4.2
7	2,438,969	1,606,649	107,729	724,592	70.3	65.9	6.3	19,614	13,867	623	5,120	73.9	70.7	4.3
8	1,353,527	914,794	16,787	421,946	68.8	67.6	1.8	19,637	13,785	605	5,244	73.3	70.2	4.2
9	1,001,420	683,754	19,117	298,549	70.2	68.3	2.7	19,651	13,834	561	5,258	73.2	70.4	3.9
10	322,597	231,418	9,959	81,220	74.8	71.7	4.1	19,667	13,865	518	5,287	73.1	70.5	3.6
11	9,495	2,941		6,554				19,687	13,899	534	5,259	73.3	70.6	3.7
12	-							19,705	13,695	526	5,480	72.2	69.5	3.7
missing	1,413	1,413												
sum	20,370,068	13,387,646	607,362	6,375,060	69.7	66.9	3.9	19,596	13,734	586	5,278	73.1	70.1	4.1

〈부표 5-13〉 2010년 월별 경제활동(KLIPS vs 통계청) - 남성

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	-							19,755	13,493	725	5,542	71.9	68.3	5.1
2	-							19,768	13,482	710	5,576	71.8	68.2	5.0
3	-							19,783	13,690	630	5,463	72.4	69.2	4.4
4	-							19,797	13,917	610	5,261	73.4	70.3	4.2
5	-							19,819	14,111	512	5,185	73.8	71.2	3.5
6	-							19,838	14,065	571	5,202	73.8	70.9	3.9
7	1,361,332	847,766	45,694	467,872	65.6	62.3	5.1	19,862	14,122	603	5,149	74.1	71.1	4.1
8	6,764,160	4,461,173	127,518	2,175,468	67.8	66.0	2.8	19,879	14,015	523	5,351	73.1	70.5	3.6
9	3,773,990	2,532,565	85,451	1,155,974	69.4	67.1	3.3	19,891	14,003	523	5,377	73.0	70.4	3.6
10	6,117,132	4,217,842	97,348	1,801,941	70.5	69.0	2.3	19,907	14,094	511	5,303	73.4	70.8	3.5
11	1,560,983	1,175,752	15,899	369,331	76.3	75.3	1.3	19,927	14,068	480	5,370	73.1	70.6	3.3
12	222,820	158,351		64,469				19,956	13,929	535	5,495	72.5	69.8	3.7
missing	-													
sum	19,800,416	13,393,450	371,910	6,035,055	69.9	67.9	2.9	19,849	13,916	578	5,356	73.0	70.1	4.0

〈부표 5-14〉 2011년 월별 경제활동(KLIPS vs 통계청) - 남성

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	11,717	3,228		8,489				19,976	13,664	555	5,742	71.3	68.4	3.9
2	54,966	36,854	2,410	15,702	71.4	67.0	6.1	19,991	13,774	649	5,570	72.1	68.9	4.5
3	-							20,010	14,007	645	5,351	73.3	70.0	4.4
4	-							20,025	14,158	589	5,289	73.6	70.7	4.0
5	-							20,054	14,278	503	5,259	73.8	71.2	3.4
6	-							20,063	14,345	505	5,221	74.0	71.5	3.4
7	394,527	239,886	15,620	139,022	64.8	60.8	6.1	20,092	14,285	534	5,271	73.8	71.1	3.6
8	8,007,960	5,184,011	242,403	2,581,546	67.8	64.7	4.5	20,108	14,277	487	5,350	73.4	71.0	3.3
9	4,274,212	2,864,235	100,046	1,309,932	69.4	67.0	3.4	20,114	14,180	484	5,457	72.9	70.5	3.3
10	4,329,164	3,092,355	82,731	1,154,078	73.3	71.4	2.6	20,126	14,330	474	5,321	73.6	71.2	3.2
11	3,550,479	2,658,118	39,041	853,320	76.0	74.9	1.4	20,149	14,346	474	5,341	73.5	71.2	3.2
12	282,055	203,230	4,420	74,404	73.6	72.1	2.1	20,198	14,179	469	5,540	72.6	70.2	3.2
missing	-													
sum	20,905,080	14,281,917	486,670	6,136,492	70.9	68.3	3.8	20,076	14,152	531	5,393	73.2	70.5	3.6

〈부표 5-15〉 2012년 월별 경제활동(KLIPS vs 통계청) - 남성

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	-							20,213	13,967	521	5,728	71.7	69.1	3.6
2	-							20,235	14,023	599	5,619	72.2	69.3	4.1
3	-							20,255	14,199	576	5,489	72.9	70.1	3.9
4	-							20,271	14,392	553	5,326	73.7	71.0	3.7
5	-							20,303	14,557	482	5,253	74.1	71.7	3.2
6	107,791	64,144	5,998	37,649	65.1	59.5	8.6	20,316	14,546	512	5,268	74.1	71.6	3.4
7	6,124,628	3,840,416	176,999	2,107,213	65.6	62.7	4.4	20,349	14,570	497	5,280	74.1	71.6	3.3
8	6,379,380	4,414,941	155,784	1,808,655	71.6	69.2	3.4	20,364	14,499	479	5,393	73.5	71.2	3.2
9	4,237,579	3,087,036	74,960	1,075,583	74.6	72.8	2.4	20,375	14,527	465	5,375	73.6	71.3	3.1
10	2,354,940	1,683,953	47,549	623,438	73.5	71.5	2.7	20,388	14,598	436	5,360	73.7	71.6	2.9
11	1,474,241	1,048,448	28,708	397,085	73.1	71.1	2.7	20,411	14,512	449	5,449	73.3	71.1	3.0
12	438,549	299,812		138,737				20,463	14,283	487	5,704	72.1	69.8	3.3
missing	581	581												
sum	21,117,689	14,439,330	489,998	6,188,361	70.6	67.8	4.0	20,329	14,390	505	5,437	73.3	70.8	3.4

〈부표 5-16〉 2013년 월별 경제활동(KLIPS vs 통계청) - 남성

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	-							20,472	14,146	528	5,807	71.6	69.1	3.6
2	-							20,500	14,063	616	5,825	71.6	68.6	4.2
3	-							20,514	14,339	535	5,642	72.5	69.9	3.6
4	-							20,534	14,579	498	5,446	73.5	71.0	3.3
5	-							20,557	14,719	486	5,362	73.9	71.6	3.2
6	1,212,936	720,936	55,347	436,653	64.0	59.4	7.1	20,585	14,739	503	5,350	74.0	71.6	3.3
7	5,697,231	3,776,228	121,779	1,799,225	68.4	66.3	3.1	20,609	14,735	534	5,347	74.1	71.5	3.5
8	4,480,257	2,906,429	82,209	1,491,619	66.7	64.9	2.8	20,620	14,661	500	5,458	73.5	71.1	3.3
9	4,143,698	2,864,413	86,280	1,193,006	71.2	69.1	2.9	20,632	14,731	456	5,447	73.6	71.4	3.0
10	3,839,617	2,869,615	67,453	902,549	76.5	74.7	2.3	20,644	14,781	457	5,416	73.8	71.6	3.0
11	1,332,794	985,795	26,648	320,350	76.0	74.0	2.6	20,659	14,792	426	5,428	73.7	71.6	2.8
12	606,031	429,696	15,514	160,822	73.5	70.9	3.5	20,673	14,595	467	5,619	72.8	70.6	3.1
missing	-													
sum	21,312,564	14,553,110	455,231	6,304,223	70.9	68.5	3.5	20,583	14,573	501	5,512	73.2	70.8	3.3

〈부표 5-17〉 2014년 월별 경제활동(KLIPS vs 통계청) - 남성

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	-							20,688	14,523	527	5,639	72.7	70.2	3.5
2	-							20,709	14,517	652	5,538	73.3	70.1	4.3
3	3,031,172	1,890,313	68,205	1,072,654	64.6	62.4	3.5	20,726	14,633	609	5,495	73.5	70.6	4.0
4	6,565,338	4,558,548	122,672	1,884,119	71.3	69.4	2.6	20,743	14,873	603	5,272	74.6	71.7	3.9
5	6,305,651	4,512,622	120,633	1,672,396	73.5	71.6	2.6	20,764	14,909	525	5,319	74.4	71.8	3.4
6	4,000,831	2,897,037	92,566	1,011,228	74.7	72.4	3.1	20,783	14,922	541	5,312	74.4	71.8	3.5
7	978,117	676,408	3,047	298,661	69.5	69.2	0.4	20,802	15,019	529	5,255	74.7	72.2	3.4
8	341,002	198,970	944	141,087	58.6	58.3	0.5	20,827	14,954	542	5,330	74.4	71.8	3.5
9	257,710	172,133	3,021	82,555	68.0	66.8	1.7	20,847	14,947	526	5,368	74.3	71.7	3.4
10	13,728	13,728						20,863	14,980	511	5,369	74.3	71.8	3.3
11	-							20,886	14,996	496	5,397	74.2	71.8	3.2
12	-							20,907	14,760	520	5,614	73.1	70.6	3.4
missing	-													
sum	21,493,549	14,919,760	411,089	6,162,700	68.6	67.2	2.1	20,795	14,836	549	5,409	74.0	71.3	3.6

〈부표 5-18〉 2015년 월별 경제활동(KLIPS vs 통계청) - 남성

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	-							20,931	14,631	578	5,719	72.7	69.9	3.8
2	-							20,959	14,671	675	5,629	73.1	70.0	4.4
3	501	501						20,983	14,793	616	5,574	73.4	70.5	4.0
4	993,815	554,661	29,934	409,220	58.8	55.8	5.1	21,010	15,001	592	5,418	74.2	71.4	3.8
5	5,507,883	3,758,282	75,399	1,674,202	69.6	68.2	2.0	21,035	15,103	580	5,357	74.5	71.8	3.7
6	4,068,926	2,772,792	58,081	1,238,053	69.6	68.1	2.1	21,061	15,059	611	5,382	74.4	71.5	3.9
7	5,267,231	3,695,772	94,808	1,476,651	72.0	70.2	2.5	21,079	15,135	598	5,343	74.7	71.8	3.8
8	3,641,693	2,529,915	52,208	1,059,570	70.9	69.5	2.0	21,098	15,064	546	5,490	74.0	71.4	3.5
9	827,724	575,199	8,631	243,894	70.5	69.5	1.5	21,118	15,078	531	5,512	73.9	71.4	3.4
10	815,944	580,034	16,434	219,476	73.1	71.1	2.8	21,136	15,112	499	5,527	73.9	71.5	3.2
11	539,970	402,717	4,308	132,946	75.4	74.6	1.1	21,154	15,083	514	5,567	73.7	71.3	3.3
12	-							21,175	14,950	510	5,709	73.0	70.6	3.3
missing	-													
sum	21,663,687	14,869,873	339,801	6,454,013	70.0	68.4	2.4	21,062	14,973	571	5,519	73.8	71.1	3.7

〈부표 5-19〉 2016년 월별 경제활동(KLIPS vs 통계청) - 남성

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	-							21,185	14,766	583	5,843	72.4	69.7	3.8
2	-							21,205	14,716	742	5,742	72.9	69.4	4.8
3	3,445	3,445						21,229	14,924	703	5,609	73.6	70.3	4.5
4	2,495,746	1,538,136	53,499	904,112	63.8	61.6	3.4	21,248	15,065	628	5,551	73.9	70.9	4.0
5	6,666,607	4,471,873	94,824	2,099,911	68.5	67.1	2.1	21,271	15,209	600	5,472	74.3	71.5	3.8
6	4,919,590	3,535,749	80,623	1,303,218	73.5	71.9	2.2	21,289	15,264	603	5,432	74.5	71.7	3.8
7	4,185,816	2,816,464	92,981	1,276,371	69.5	67.3	3.2	21,298	15,292	604	5,396	74.7	71.8	3.8
8	2,057,409	1,459,019	66,633	531,757	74.2	70.9	4.4	21,309	15,257	603	5,439	74.5	71.6	3.8
9	558,532	349,459	15,378	193,696	65.3	62.6	4.2	21,320	15,244	586	5,482	74.3	71.5	3.7
10	616,370	473,455	4,145	138,770	77.5	76.8	0.9	21,332	15,295	538	5,505	74.2	71.7	3.4
11	313,216	233,485	2,876	76,855	75.5	74.5	1.2	21,344	15,304	489	5,555	74.0	71.7	3.1
12	-							21,356	15,120	484	5,746	73.1	70.8	3.1
missing	-													
sum	21,816,732	14,881,085	410,958	6,524,689	71.0	69.1	2.7	21,282	15,121	597	5,564	73.9	71.1	3.8

〈부표 5-20〉 2009년 월별 경제활동(KLIPS vs 통계청) - 여성

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	-							20,380	9,456	282	10,640	47.8	46.4	2.9
2	-							20,399	9,384	320	10,707	47.5	46.0	3.3
3	38,855	35,861	-	2,993				20,419	9,577	306	10,540	48.4	46.9	3.1
4	5,824,343	2,366,637	112,695	3,345,012	42.6	40.6	4.5	20,440	9,791	323	10,333	49.4	47.9	3.2
5	6,679,834	2,908,987	159,327	3,611,521	45.9	43.5	5.2	20,464	9,864	326	10,261	49.9	48.2	3.2
6	2,933,423	1,286,342	50,377	1,596,703	45.6	43.9	3.8	20,489	9,999	342	10,139	50.5	48.8	3.3
7	2,280,318	1,055,193	59,208	1,165,918	48.9	46.3	5.3	20,511	9,948	308	10,249	50.0	48.5	3.0
8	1,451,753	566,258	23,935	861,560	40.7	39.0	4.1	20,531	9,834	304	10,399	49.3	47.9	3.0
9	1,029,679	410,400	22,454	596,825	42.0	39.9	5.2	20,550	9,967	266	10,314	49.8	48.5	2.6
10	316,812	140,534	10,775	165,503	47.8	44.4	7.1	20,570	9,997	277	10,295	50.0	48.6	2.7
11	19,346		5,202	14,144				20,590	9,924	286	10,392	49.5	48.2	2.8
12	-							20,611	9,522	315	10,773	47.7	46.2	3.2
missing	5,537	5,537	-	-										
sum	20,579,900	8,775,749	443,972	11,360,179	44.3	42.5	4.7	20,496	9,772	305	10,420	49.2	47.7	3.0

〈부표 5-21〉 2010년 월별 경제활동(KLIPS vs 통계청) - 여성

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	-							20,632	9,367	493	10,763	47.8	45.4	5.0
2	-							20,652	9,376	463	10,808	47.7	45.4	4.7
3	-							20,672	9,695	372	10,611	48.7	46.9	3.7
4	-							20,693	9,995	320	10,370	49.9	48.3	3.1
5	3,249	-	-	3,249				20,714	10,191	272	10,250	50.5	49.2	2.6
6	-							20,734	10,201	316	10,213	50.7	49.2	3.0
7	1,416,464	555,635	24,590	836,240	41.0	39.2	4.2	20,753	10,190	337	10,234	50.7	49.1	3.2
8	7,045,759	3,060,431	55,532	3,929,796	44.2	43.4	1.8	20,772	9,991	309	10,464	49.6	48.1	3.0
9	4,009,962	1,772,378	48,393	2,189,191	45.4	44.2	2.7	20,790	10,062	343	10,394	50.0	48.4	3.3
10	6,008,311	2,756,796	85,304	3,166,212	47.3	45.9	3.0	20,809	10,092	312	10,408	50.0	48.5	3.0
11	1,681,429	772,597	32,630	876,201	47.9	45.9	4.1	20,828	10,039	257	10,539	49.4	48.2	2.5
12	236,496	118,119	3,3706	114,671	51.5	49.9		20,847	9,756	322	10,770	48.3	46.8	3.2
missing	1,697	1,697	-	-										
sum	20,403,367	9,037,652	246,449	11,115,560	45.2	43.7	3.1	20,741	9,913	343	10,485	49.4	47.8	3.4

〈부표 5-22〉 2011년 월별 경제활동(KLIPS vs 통계청) - 여성

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	28,402	13,357	-	15,045				20,867	9,515	356	10,988	47.3	45.6	3.6
2	16,592	8,411	-	8,181				20,887	9,566	450	10,878	47.9	45.8	4.5
3	-							20,908	9,827	421	10,649	49.1	47.0	4.1
4	-							20,928	10,150	357	10,424	50.2	48.5	3.4
5	-							20,949	10,370	310	10,264	51.0	49.5	2.9
6	-							20,970	10,422	333	10,220	51.3	49.7	3.1
7	427,787	176,971	-	248,330				20,995	10,351	309	10,343	50.7	49.3	2.9
8	8,081,163	3,424,603	101,631	4,554,929	43.6	42.4	2.9	21,011	10,232	273	10,512	50.0	48.7	2.6
9	4,369,729	1,981,561	67,415	2,320,752	46.9	45.3	3.3	21,027	10,135	281	10,608	49.5	48.2	2.7
10	4,532,572	2,118,222	44,261	2,370,089	47.7	46.7	2.0	21,043	10,332	265	10,439	50.4	49.1	2.5
11	3,425,714	1,664,087	32,396	1,729,230	49.5	48.6	1.9	21,059	10,256	263	10,548	49.9	48.7	2.5
12	221,484	129,853	3,208	88,423	60.1	58.6	2.4	21,075	9,947	286	10,854	48.5	47.2	2.8
missing	-													
sum	21,103,444	9,517,066	251,398	11,334,980	49.6	48.3	2.5	20,977	10,092	325	10,561	49.7	48.1	3.1

〈부표 5-23〉 2012년 월별 경제활동(KLIPS vs 통계청) - 여성

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동 인구	경제활동 참가율	고용률	실업률	15세 이상인구	취업자	실업자	비경제 활동 인구	경제활동 참가율	고용률	실업률
1	-							21,108	9,773	333	11,009	47.8	46.3	3.3
2	-							21,134	9,764	449	10,925	48.3	46.2	4.4
3	-							21,160	10,072	366	10,716	49.4	47.6	3.5
4	-							21,188	10,361	343	10,480	50.5	48.9	3.2
5	-							21,216	10,566	316	10,326	51.3	49.8	2.9
6	150,632	50,129	-	100,503				21,245	10,580	316	10,354	51.3	49.8	2.9
7	6,194,032	2,572,443	104,629	3,516,960	43.2	41.5	3.9	21,280	10,534	292	10,447	50.9	49.5	2.7
8	6,409,332	2,973,934	96,325	3,339,073	47.9	46.4	3.1	21,301	10,352	288	10,649	50.0	48.6	2.7
9	4,403,042	2,177,744	36,261	2,189,037	50.3	49.5	1.6	21,321	10,469	280	10,567	50.4	49.1	2.6
10	2,477,249	1,128,365	37,770	1,311,114	47.1	45.5	3.2	21,342	10,479	280	10,583	50.4	49.1	2.6
11	1,315,440	590,579	37,832	687,029	47.8	44.9	6.0	21,364	10,426	257	10,674	50.0	48.8	2.4
12	377,358	165,489	2,951	208,917	44.6	43.9	1.8	21,385	10,115	260	11,005	48.5	47.3	2.5
missing	600	-	-	600										
sum	21,327,684	9,658,682	315,769	11,353,233	46.8	45.3	3.3	21,254	10,291	315	10,645	49.9	48.4	3.0

〈부표 5-24〉 2013년 월별 경제활동(KLIPS vs 통계청) - 여성

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	-							21,405	9,911	317	11,169	47.8	46.3	3.1
2	-							21,426	9,920	381	11,128	48.1	46.3	3.7
3	-							21,447	10,166	347	10,922	49.1	47.4	3.3
4	-							21,468	10,519	325	10,628	50.5	49.0	3.0
5	-							21,490	10,681	319	10,490	51.2	49.7	2.9
6	1,408,451	669,722	19,051	719,679	48.9	47.6	2.8	21,512	10,734	310	10,456	51.4	49.9	2.8
7	5,936,263	2,527,659	67,311	3,341,294	43.7	42.6	2.6	21,535	10,746	298	10,497	51.3	49.9	2.7
8	4,604,058	2,090,798	28,916	2,484,344	46.0	45.4	1.4	21,548	10,623	284	10,636	50.6	49.3	2.6
9	3,956,153	1,912,047	57,304	1,986,803	49.8	48.3	2.9	21,561	10,737	275	10,560	51.0	49.8	2.5
10	3,662,551	1,690,595	25,231	1,946,726	46.8	46.2	1.5	21,573	10,765	276	10,533	51.2	49.9	2.5
11	1,402,985	682,307	7,077	713,601	49.1	48.6	1.0	21,586	10,728	264	10,587	51.0	49.7	2.4
12	566,521	294,551	-	271,970				21,599	10,368	310	10,917	49.5	48.0	2.9
missing	-													
sum	21,536,982	9,867,677	204,889	11,464,416	47.4	46.4	2.0	21,513	10,492	309	10,710	50.2	48.8	2.9

〈부표 5-25〉 2014년 월별 경제활동(KLIPS vs 통계청) - 여성

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	-							21,616	10,246	360	11,015	49.0	47.4	3.4
2	-							21,634	10,298	520	10,809	50.0	47.6	4.8
3	3,457,911	1,368,527	46,378	2,043,006	40.9	39.6	3.3	21,652	10,545	416	10,697	50.6	48.7	3.8
4	6,698,893	3,066,992	90,332	3,541,569	47.1	45.8	2.9	21,670	10,813	438	10,428	51.9	49.9	3.9
5	5,855,953	2,817,362	70,185	2,968,406	49.3	48.1	2.4	21,689	10,888	419	10,372	52.2	50.2	3.7
6	4,267,774	2,050,938	50,929	2,165,907	49.2	48.1	2.4	21,708	10,941	409	10,354	52.3	50.4	3.6
7	739,674	344,254	24,769	370,651	49.9	46.5	6.7	21,726	10,950	386	10,382	52.2	50.4	3.4
8	366,667	165,408	-	201,259				21,745	10,916	350	10,467	51.9	50.2	3.1
9	332,771	148,458	-	184,313				21,764	10,969	327	10,477	51.9	50.4	2.9
10	5,563	5,563	-	-				21,783	10,979	339	10,468	51.9	50.4	3.0
11	4,565	4,565	-	-				21,802	10,966	328	10,505	51.8	50.3	2.9
12	-							21,822	10,605	362	10,844	50.3	48.6	3.3
missing	1,648	1,648	-	-										
sum	21,731,418	9,973,715	282,593	11,475,110	47.3	45.6	3.5	21,718	10,760	388	10,568	51.3	49.5	3.5

〈부표 5-26〉 2015년 월별 경제활동(KLIPS vs 통계청) - 여성

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	-							21,845	10,486	403	10,962	49.8	48.0	3.7
2	-							21,868	10,540	531	10,801	50.6	48.2	4.8
3	-							21,892	10,705	458	10,723	51.0	48.9	4.1
4	1,233,963	522,296	16,778	694,889	43.7	42.3	3.1	21,915	10,914	454	10,554	51.8	49.8	4.0
5	5,754,446	2,413,237	113,831	3,227,379	43.9	41.9	4.5	21,940	11,102	438	10,406	52.6	50.6	3.8
6	4,080,832	1,906,838	30,732	2,143,261	47.5	46.7	1.6	21,964	11,136	428	10,388	52.7	50.7	3.7
7	5,301,685	2,545,756	48,736	2,707,192	48.9	48.0	1.9	21,976	11,164	405	10,408	52.6	50.8	3.5
8	3,392,332	1,675,584	66,964	1,649,784	51.4	49.4	3.8	21,988	11,082	367	10,532	52.1	50.4	3.2
9	849,183	410,132	35,176	403,875	52.4	48.3	7.9	22,001	11,199	334	10,478	52.4	50.9	2.9
10	749,355	349,598	26,026	373,731	50.1	46.7	6.9	22,014	11,183	346	10,485	52.4	50.8	3.0
11	551,964	239,409	4,998	307,557	44.3	43.4	2.0	22,027	11,168	322	10,532	52.2	50.7	2.8
12	-							22,041	10,932	350	10,760	51.2	49.6	3.1
missing	-													
sum	21,913,759	10,062,850	343,242	11,507,668	47.8	45.8	4.0	21,956	10,967	403	10,586	51.8	50.0	3.6

〈부표 5-27〉 2016년 월별 경제활동(KLIPS vs 통계청) - 여성

월	KLIPS							통계청						
	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률	15세 이상 인구	취업자	실업자	비경제 활동인구	경제활동 참가율	고용률	실업률
1	-							22,054	10,674	410	10,963	50.3	48.4	3.7
2	-							22,069	10,703	575	10,798	51.1	48.5	5.1
3	-							22,084	10,887	453	10,749	51.3	49.3	4.0
4	2,960,731	1,142,041	44,031	1,774,659	40.1	38.6	3.7	22,100	11,072	450	10,569	52.2	50.1	3.9
5	6,972,186	3,157,259	70,090	3,744,837	46.3	45.3	2.2	22,116	11,235	408	10,460	52.7	50.8	3.5
6	4,865,397	2,453,196	41,315	2,370,886	51.3	50.4	1.7	22,133	11,310	398	10,427	52.9	51.1	3.4
7	3,931,350	1,924,894	66,274	1,940,182	50.6	49.0	3.3	22,144	11,293	374	10,468	52.7	51.0	3.2
8	1,814,341	877,782	26,603	909,956	49.8	48.4	2.9	22,156	11,255	385	10,501	52.6	50.8	3.3
9	524,868	237,801	10,350	276,717	47.3	45.3	4.2	22,168	11,284	397	10,490	52.7	50.9	3.4
10	702,649	391,709	7,528	303,412	56.8	55.7	1.9	22,180	11,290	385	10,508	52.6	50.9	3.3
11	314,019	151,387	1,774	160,857	48.8	48.2	1.2	22,193	11,296	361	10,536	52.5	50.9	3.1
12	-							22,205	11,036	377	10,781	51.4	49.7	3.3
missing	-													
sum	22,085,540	10,336,069	267,966	11,481,506	48.9	47.6	2.6	22,134	11,111	414	10,604	52.1	50.2	3.6

제 6 장
한국노동패널조사에 웹서베이 도입 가능성 검토를 위한 연구

제1절 서 론

본 연구는 한국노동패널조사에 웹서베이 도입 가능성을 검토하고자 한다. 웹서베이는 우편 조사처럼 누군가의 도움을 받지 않고 자기기입식으로 진행되는 설문조사의 한 종류라는 점에서 면접조사와 구별되며, 조사의 매개체로 종이 설문지가 아니라 웹 설문지가 이용된다는 점에서 다른 자기기입식 조사와 구별된다.

우리나라 대규모 패널조사에서 웹서베이는 도입 경험이 없으므로, 본 연구에서는 왜 웹서베이 도입 논의가 필요한지부터 검토한다. 조사관련 학계, 업계에서는 이미 2000년대 이래 웹서베이가 다양한 영역에서 실시되어 왔으며, 웹서베이가 갖는 장단점에 대해 잘 정리가 되어 있다. 본 연구는 도입 가능성 검토를 위해 먼저 웹서베이가 갖는 장단점을 문헌연구를 통해 요약한다.

다음으로 대규모 패널조사에서 웹서베이를 도입하거나 도입을 위해 실험이 이루어지고 있는 사례들을 검토한다. 이를 통해 한국노동패널조사에 웹서베이가 도입될 가능성이 있는지, 도입을 한다면 어느 정도의 규모로, 어떤 식의 도입과정을 거치는 것이 합리적일지 생각해 보고자 한다.

여기에 덧붙여 웹서베이 도입을 위한 사전 단계로서 파일럿 조사에 대한 제안을 한다. 파일럿 조사 결과는 이 보고서 제7장에 수록된다.

제2절 웹서베이의 장단점

Smyth & Pearson(2011)에 따르면 웹서베이의 장점은 다음과 같다. 첫째, 비용이 적게 든다. 웹서베이의 비용은 주로 웹서베이를 위한 하드웨어 및 소프트웨어 구축 비용이며, 일단 구축이 되고 나면 조사대상자에게 비용을 지급하는 정도를 빼고는 추가적인 비용이 들지 않는다는 장점이 있다.

둘째, 좀 더 빠르게 조사를 진행할 수 있다. 웹서베이에 응하게만 만들 수 있다면, 면접원이 이동하는 시간이나 우편 설문지가 닿는 시간 같은 추가적인 시간이 필요하지 않다. 게다가 조사 결과가 웹에 입력되는 순간 이미 데이터로 존재하기 때문에 추가로 펀칭을 해야 하는 시간과 수고, 펀칭과정에서 나타날 수 있는 오류 가능성도 낮출 수 있다.

셋째, 문항의 스킵(skip) 구조가 설문조사 응답을 어렵게 만드는 경우가 많은데, 웹서베이는 이미 프로그램 된 문항이 자동으로 뜨기 때문에 오류 가능성이 낮다. 필요하다면, 프로그램상으로 임의 문항 할당 등이 자동으로 가능하기 때문에 설문조사에서 유연성을 확보할 수 있다는 장점도 있다.

그러나 웹서베이는 커버리지(coverage) 측면에서 중대한 약점이 있다. 인터넷 접근성이 높지 않은 계층은 웹서베이 방식으로 조사하기 어렵다. 이를테면 미국에서 65세 이상은 35%만 인터넷 접근성을 가지고, 비히스패닉 흑인은 60% 미만이 접근성을 갖고 있다는 연구가 있다.

또한 다른 조사에 비해 응답률도 낮다고 한다. 메타 분석에 따르면 웹서베이는 30% 정도의 응답률을 보여 우편조사에 비해서도 10% 이상 낮다고 한다(de Leeuw & Hox, 2011).

측정(measurment)의 측면에서 웹서베이는 몇 가지 장점이 있다. 우선,

응답자가 민감하게 생각할 수 있는 질문에서 자기기입식이라는 특징 때문에 프라이버시 보호가 가능하고, 이로 인해 좀 더 정직한 응답을 유도할 수 있다는 장점이 있다.

또 다른 큰 장점은 설문구성에서 음악, 비디오, 그림을 넣을 수 있어 질문의 한계를 넓혀준다는 점이다. 다만, 잘못 설계될 경우 5점 척도가 좋으냐 7점 척도가 좋으냐, 스케일을 표현하는 단어를 무엇으로 선택하느냐에 따라 응답결과가 달라지듯 오류를 야기할 가능성이 있다는 점은 주의해야 할 사항이다.

설문에 대한 적극성을 가진 응답자라면 설문의 속도를 스스로 통제하면서 좀 더 철저한 답변을 기대할 수 있다는 것도 장점이다. 그러나 웹서베이를 하면서 채팅을 하고 웹서핑을 하는 등 서베이에 대한 집중력을 해치는 행동이 함께 발생할 가능성이 높다는 단점도 동시에 존재한다.

이 외에도 면접원이 없으므로 일단 설문을 시작한다 하더라도 설문을 마쳐야 한다는 압박에서 자유로워 완성된 설문결과를 확보할 가능성이 낮을 수 있다는 점도 단점으로 지적될 수 있다.

제3절 웹서베이의 단점을 극복할 대안으로서 혼합조사(mixed mode)

이상에서 본 여러 단점 중에 노동패널과 같은 대규모 설문조사에서 가장 중요하게 부각되는 단점은 커버리지와 낮은 응답률 문제일 것이다. de Leeuw & Hox(2011)는 웹서베이가 갖는 낮은 커버리지와 낮은 응답률을 감안하면, 웹서베이 단독조사가 아니라 웹서베이를 다른 조사방법과 혼합하는 혼합조사(mixed mode) 방식이 대안이 될 수 있다고 하였다. 이렇게 하면 웹서베이의 장점인 비용절감을 어느 정도 추구하면서도 웹서베이의 단점인 낮은 커버리지 문제와 낮은 응답률 문제를 우회할 수 있다. 물론 이는 다른 조사방법으로 조사된 데이터의 비교가능성 문제를 야기할 수 있다.

어떻게 혼합조사를 할 것인가에 대해서도 다양한 조합을 생각해 볼 수 있다. de Leeuw & Hox(2011)는 몇 가지 조합의 혼합조사 방법을 제시하고 있다.

첫째, 다른 조사 방식으로 컨택하고 응답방식은 하나의 방식으로 하는 혼합조사 방식이다. 응답방식이 웹서베이일 때, 커버리지 제약을 우회하는 수단으로 이러한 형태의 혼합조사 방식이 활용될 수 있다. 또는, 조사 참여를 설득할 때 좀 더 나은 다른 수단이 있을 수 있음을 감안하는 것일 수도 있다. 이를테면, 어떤 사람들은 집에 찾아가 설득하는 것이 나을 수 있고, 어떤 사람들은 집에 거의 있지 않아 이메일이나 전화를 이용하는 것이 나을 수 있다.

둘째, 설문의 일부분만 다른 조사방식을 활용하는 것이다. 이를테면, 전체조사를 CAPI(Computer-assisted personal interviewing) 같은 대면면접조사 방식으로 진행하되, 민감한 설문만 웹서베이로 진행하는 방식이 예가 될 수 있다. 이는 사회적 바람직함 편향(social desirability)을 감소시키는 방법으로 선택될 수 있겠다. 이런 방법은 각 조사방법이 가지고 있는 강점을 취하기 위해 선택된다. 또한, 조사 방식이 설문마다 다르지, 같은 설문에 대해 다르지는 않으므로 데이터 품질에서 mode effect도 발생하지 않는다는 장점이 있다.

셋째, 응답자 특성에 따라 다른 조사 방식을 제공하는 방법도 있다. 이를테면, 웹 접촉 빈도가 높은 특성을 갖는 사람들에게는 웹서베이를, 웹 접촉 빈도가 낮은 특성을 갖는 사람들에게는 CAPI를 하는 것이 예가 될 수 있다. 이는 웹서베이가 갖는 커버리지 문제를 우회하면서 각 조사방법이 가지고 있는 무응답 문제를 개선하고자 할 때 활용될 수 있다. 또 다른 사례로 전체 표본을 웹서베이로 조사하되, 웹서베이에 무응답한 사람을 CAPI나 CATI(Computer-assisted telephone interviewing)로 접촉하는 방법이다. 이는 아래에서 보겠지만 웹서베이를 도입한 대규모 패널조사에서 나타나는 혼합조사 방법이다. 다만, 이와 같이 도입할 경우 웹서베이의 비용절감 효과를 부분적으로 누릴 수 있지만, 같은 설문을 다른 조사방법으로 조사한다는 문제 때문에 데이터에서 mode effect가 나타날 개연성을 안고 있다.

넷째, 패널 조사라면 응답자 설득이 매우 중요한 첫 번째 조사에서는 대면조사를 하고, 두 번째 조사부터는 웹서베이를 하는 것도 대안이 될 수 있다. 다만, 이 경우 시간에 따른 mode effect가 나타날 개연성이 있다.

제4절 조사된 데이터에서 나타날 수 있는 이슈와 대안

조사 방식이 다를 때 조사된 결과가 다를 수 있다. 이에는 세 가지 이유가 있다(de Leeuw & Hox, 2011). 첫 번째, 각 조사방식에 잘 반응하는 표본들이 서로 다른 특성을 가진 표본들일 경우에 조사결과가 달라질 수 있다. 두 번째, 조사 방법마다 질문 포맷이 달라서 다른 응답결과가 야기될 수 있다. 세 번째, 조사 방법 자체에 내재된 편향이 다른 응답으로 이끌 수 있다.

이 중 첫 번째는 혼합조사 방법의 단점이 아니라 장점이다. 혼합조사의 원래적 목적인 커버리지 문제와 무응답 편향이 교정된 결과라 받아들여야 하기 때문이다.

두 번째와 세 번째가 문제이다. 두 번째 문제는 조사 방식마다 잘 조사되는 설문지 구성이 다를 수 있고, 개별 조사 방식에 최적화된 방식으로 설문지가 만들어졌을 경우 발생한다. 이를테면, 종이설문에서는 5점 척도가 단순히 세로로 나열되고, 웹서베이에서는 예쁘게 수직선 상에 칼라로 5점 척도가 표현되는 것이 사례가 될 수 있다. 이러한 두 번째 문제는 “unified mode design(=uni-mode design)”을 통해 회피될 수 있다. 이는 조사 방법에 관계없이 동일한 질문, 동일한 선택지를 제공하는 설문지 설계원리를 뜻한다.

세 번째 문제는 순수한 의미의 mode effect이다. 이를테면 면접조사에서 흔히 나타나는 사회적 바람직함 편향이 그런 예가 된다. 기존 연구들의 메타분석결과에 따르면, 사회적 바람직함 편향은 웹서베이가 확실히 덜하고, 오픈형 문항에 대한 응답은 좀 더 잘 나오는 경향이 있으며, 항목무응답 유발 여부는 조사마다 엇갈려 확실한 결론이 없는 상태라고 한다.

가장 분명한 차이는 사회적 바람직함 편향이므로, CAWI와 CAPI를 이용한 혼합조사라면 사회적으로 민감한 문항은 별도의 설문지를 제공하든지, 면접원이나 가족 같은 다른 응답자가 보지 못하는 곳에서 응답자가 직접 입력하도록 하는 것이 대안이 될 수 있다.

제5절 대규모 패널조사 웹서베이 도입 사례

1. The Age 55 Survey of the National Child Development Study(NCDS)[41]

이 조사는 1958년 이래 7세인 사람을 대상으로 7, 11, 16, 23, 33, 42, 44, 46, 50, 55세 당시 재접촉하여 조사하는 설계를 가지고 있는 개인 표본조사이다. 여기서 웹서베이는 웹서베이로 우선 접촉하고 안 되면 전화 조사로 넘어가는 순서형 혼합조사(sequential mixed mode) 설계로 도입되었다.

55세 조사 이전에는 대면면접조사가 일차적인 조사 방식이었고, 46세 당시 조사만 전화 조사되었다고 한다. 이 조사에서 웹서베이를 하게 된 핵심 이유는 비용절감이었다. 이 외에도 웹을 사용하면 조사대상자들이 설문을 하는 또 다른 수단을 가지게 되므로 좀 더 유연성이 확보된다는 장점, 조사 기간 단축 가능성, 좀 더 비주얼하게 서베이 할 수 있는 잠재력 등의 장점이 있을 것으로 판단되었다고 한다.

이 조사가 웹서베이를 도입하면서 세 조사의 사례를 참조했다고 한다. 첫째는, “the Longitudinal Study of Young People in England”이다. 이 조사는 웹서베이를 활용한 결과 5차 조사에서 응답자의 37%, 7차 조사에서는 46%까지 웹을 이용한 응답률이 상승했다고 한다. 이를 보고 조사비 절감을 위해 웹서베이의 가능성이 있다고 판단했다고 한다.

또한, 미국 “Health and Retirement Study(HRS)”도 참고했다고 한다.

41) 이하의 내용은 “Technical Report of the 1958 National Child Development Study: Age 55 survey(2013/2014)”를 정리한 것이다.

여기서는 조사와 조사 사이 공백 기간("between wave")에 웹을 이용해 간단한 설문을 응답하도록 응답자를 접촉했다. 50세 이상 고령층이 응답 대상이었고, 70~75%의 응답률을 얻었다고 한다.

이 외에도 영국의 "the British Crime Survey"도 참고하였는데, 실험연구를 통해 연령이 증가함에 따라 온라인 응답이 증가하는 결과를 얻었다고 한다. 따라서 고령자를 대상으로 하는 이 연구에서도 웹서베이 가능성이 있겠다고 판단했다고 한다.

NCDS는 실제 웹서베이를 탑재하기 전에 비패널표본을 만들어 테스트를 거쳤다. 또한, 엄격한 평가를 위해 웹서베이를 본조사에 탑재하기 전에 1,500패널표본을 상대로 응답률 및 설문에 제대로 응답하는지를 평가하는 실험연구를 했다. 전화면접조사 대 웹서베이-전화면접 혼합조사로 실험연구를 했다. 실험연구 결과 혼합조사 대상으로 할당된 표본 중 62%가 온라인으로 완성했고, 21%가 전화면접까지 가서 총 83%의 응답률을 기록했다고 한다.

설문구성 시 soft checks(문제가 있어 보이는 응답 시 재질문)와 hard checks(문제가 있어 보이는 응답 시 수정될 때까지 재질문하고 다음으로 넘어가지 않는 것)가 똑같이 설계되게 노력했다. 다만, 이런 체크가 많으면, 면접원 없이 응답하는 웹서베이에서는 응답에 혼란을 겪을 수 있다. 따라서 이런 체크를 최소화하고, 특히 소프트 체크는 설문형태를 띠게 변경하였다. "앞선 무슨 무슨 문항에서 이렇게 응답하셨는데 이 문항에 이렇게 대답한 것이 정확합니까? 네, 정확합니다/ 아니오, 대답을 변경하겠습니다" 형식으로 질문했다. 또한, 하드체크가 필요한 상황이 되면 에러 메시지가 뜨게 했다.

과년도에 응답한 정보와 상호 체크하는 문항이 있는 경우 이전 응답이 맞는지 틀리는지 질문해 대답하게 하고, 틀리다고 응답하면 수정하게 해서 과거 정보를 클리닝하는 방식을 선택했다고 한다.

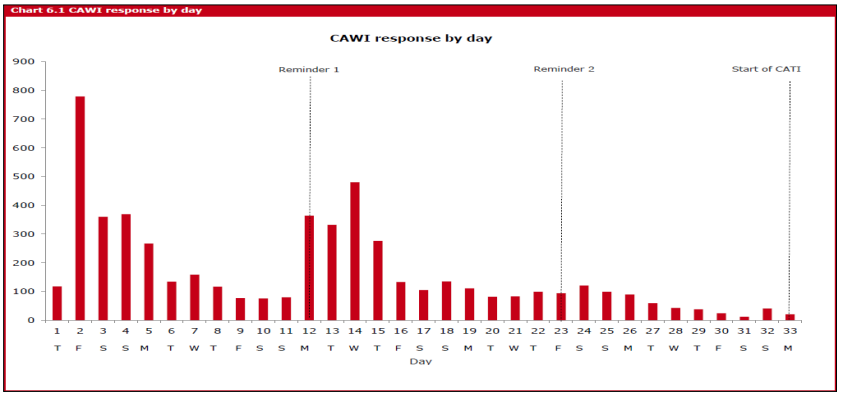
30분 분량의 설문으로 구성하였으며, 모든 표본은 온라인으로 응답하도록 가이드되었다고 한다. 그러나, 5주의 조사기간 안에 두 번의 편지/이메일을 통한 고지 이후에도 응답하지 않으면 전화면접 단계로 넘어갔다. 실험을 위해 전체 10,558표본 중 임의추출된 1,499표본은 전화면접만 수

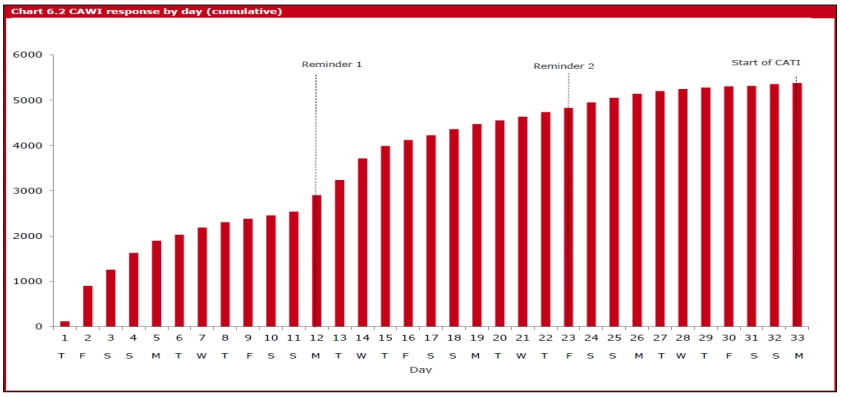
행했다.

CAWI 접촉 단계에서는 알림 편지, 이메일, 서베이 리플렛을 보냈다고 한다. 먼저 온라인으로 응해 줄 것을 청하고, 온라인이 불가능할 시 몇 주 뒤에 전화면접이 이루어질 것임을 공지하였다. 2013년 9월 12일 온라인으로 응해 줄 것을 청하는 첫 공지를 보냈고, 23일 재알림 메일을 발송했다. 10월 23일 재재알림 메일을 발송했고, 첫 알림 공지 5주 후 전화면접단계로 넘어갔다.

총 8,958표본을 완성해 응답률 78%를 기록했고, 주소파악 등 유효표본 중 88%가 성공했다. 총 성공표본 중 65%는 웹서베이로 완성했으며, 35%는 전화면접으로 완성했다. 재알림과 응답률의 관계는 [그림 6-1]을 참조하면 되겠다.

〔그림 6-1〕 NCDS의 CAWI 응답률 추이





웹서베이 응답자 중 48%가 노트북, 32%는 데스크탑, 13%는 태블릿, 1%는 스마트폰으로 수행했다고 한다. 전화면접보다 웹서베이가 8분 정도 더 걸렸다고 보고되었다. 평균 각각 27분 대 35분 걸렸다고 한다.

웹서베이→전화서베이로 넘어간 순서형 복수조사도구 방법의 응답률은 83%, 전화면접만 한 경우 응답률은 78%로 복수조사도구 방법의 응답이 더 높게 나타났다고 한다.

2. Health and Retirement Study(HRS)[42]

미국 고령자를 대상으로 하는 이 서베이는 1992년 이래 50세 이상 격년 조사하는 패널조사로 대면면접과 전화면접을 병행하여 조사되어 왔다.

2006년 이래 표본의 절반씩 강화된 대면면접이라는 이름하에 물리적 힘이나 생물학적 정보 등의 수집을 포함하여 조사하는 특징이 있으며, 나머지 절반은 전화 면접하였다.

2018년부터는 전화면접 대상에겐 웹서베이를 제공할 계획을 가지고 있으며, 이를 위해 2015년 봄 650명의 참가자를 대상으로 파일럿 서베이를 실시하였다.

설문 전체가 아니라 코어 질문만을 대상으로 파일럿 서베이를 실시하였다. 인구학적 배경, 건강, 건강보험과 서비스, 소득, 자산 등 약 40% 분량의 설문이다. 이 외에도 소셜 미디어 사용, 건강보험 개혁, 자원봉사와 돌봄활동 등 3개의 새로운 토픽이 웹서베이로 테스트되었다.

HRS는 웹서베이 이행과정에서 설문지가 웹서베이에 적합한지 검토하는 것에 주력했다고 한다. 면접조사→자기기입식으로 바뀌는 것이기 때문에 설문이 자기기입식에 적합한지 검토할 필요가 있었기 때문이다.

또한, 웹서베이가 응답자에게 끔찍한 경험이 되어서는 안 되므로, 좀 더 명확하게 이해되도록 질문지를 개선하고, 설문문항 간 네비게이션이 쉽게 되도록 개선했으며, 디자인 요소도 첨가하였다.

---

42) 이하의 내용은 Ofstedal, Mary Beth, Mick P. Couper, and Colleen McClain(2016), “Transitioning a Panel Survey to Web: Experiences from the Health and Retirement Study”, Working Paper를 정리한 내용이다.

설문 개선 과정에서 응답의 결측(missing value) 허용범위, “모른다” 응답 허용범위가 고민이었다고 한다. 수치 입력의 경우 적절한 범위가 아니면 하드체크나 소프트체크가 뜨도록 하였다.

얼마나 다양한 브라우저와 디바이스에서 웹서베이가 가능하도록 할 것인가를 고민하였으며, 잘 모를 경우 헬프 파일을 이용가능하게 하였고, 면접원 가이드로 활용되던 것을 프로그램에 탑재해 잘 모르는 사항을 쉽게

〈표 6-1〉 HRS의 파일럿 서베이 응답률

Table 3. Response Rates by Sample Indicators

	Percent (SE), n
Experimental Group[1]	
Non-treatment, initial sample release (n=50)	36% (6.8%), n=18
Non-treatment, no email address for R (n=72)	62.5% (5.7%), n=45
1A: Combined mail/email, email/letter same day (n=155)	66.5% (3.8%), n=103
1B: Combined mail/email, letter preceded email (n=138)	83.3% (3.2%), n=115
2: Mail only (among Rs with email addresses) (n=240)	74.2% (2.8%), n=178
	X^2=10.898**
Age	
< 60 (n=155)	63.9% (3.9%), n=99
60-69 (n=210)	77.1% (2.9%), n=162
70-79 (n=224)	70.5% (3.0%), n=158
80+ (n=66)	60.6% (6.0%), n=40
	X^2=10.692*
Gender	
Female (n=374)	69.3% (2.4%), n=259
Male (n=281)	71.2% (2.7%), n=200
	X^2=10.692*
Race/Ethnicity	
Hispanic (n=39)	51.3% (8.0%), n=20
Non-Hispanic Black (n=66)	54.5% (6.1%). n=36
Non-Hispanic White, Other, Not Ascertained (n=550)	73.3% (1.9%). n=403
	X^2=16.841***
Education[2]	
<= 12 years (n=194)	66.5% (3.4%), n=129
13-15 years (n=181)	66.9% (3.5%), n=121
16 years (n=128)	75.8% (3.8%), n=97
17+ years (n=152)	73.7% (3.6%), n=112
	X^2=5.015

주 : *** p≤.001, ** p≤.01, * p≤.05, † p≤1

자료 : Ofstedal 등(2016).

확인가능하도록 하였다. 화면에 포함되는 용어 정의 같은 것이 얼마나 이해하기 쉬운가를 평가해 이해하기 쉽게 교정하고 고치는 작업이 있었다.

스크롤링 없이도 선택지가 전부 한 화면에 보이도록 하는 화면 레이아웃의 구성 및 선택지가 너무 길 경우 선택지를 줄이는 설문 개편작업도 병행하였다고 한다. 같은 섹션의 질문이면 배너나 헤딩으로 같은 섹션에 있음을 표시하게 해주었다.

655개 표본 중 459개 표본이 설문을 부분적으로든 완전히든 웹서베이를 이용해 응답하여 70% 응답률을 보였다. 추가 5%는 로그인은 했다고 한다(34표본). 25%는 로그인조차 안 했다고 한다.

웹서베이 응답자의 약 3/4 정도는 일회 접속으로 전체 설문을 끝냈으며, 약 13% 정도는 세 번 이상 접속해서 끝냈다. 최대 접속자는 57회였다.

설문의 복잡성 때문에 초대 메일에서 가급적 노트북이나 데스크탑으로 접속해 달라고 하였다. 모바일 디바이스는 이용하지 말아달라고 하여 태블릿이나 스마트폰으로 설문을 한 사람은 8% 선에 그쳤다고 한다.

HRS는 고령자 커플당 인터뷰 유형에 따라 50~80달러를 선지급한다(prepaid). 거절 표본가구를 참여하게 만들기 위해 조사 최종 단계에서 100달러를 인센티브로 지급한다. 실험은 일단 조사에 응하게 하기 위해 웹서베이 초대 편지와 초대 이메일을 보내는 그룹1과 초대편지만 보내는 그룹2로 구분하였다. 임의표집하여 구분하였다. 초대편지만 받는 그룹2는 이후 재촉 메일도 편지로만 받게 하였다. 또한, 초대편지와 초대이메일을 모두 받는 그룹은 웹서베이 초대 편지를 발송하면서 초대 이메일을 같이 보낸 그룹(1A), 이메일 초대 3일 전 웹서베이 초대 편지를 발송한 그룹(1B)으로 구분하였다. 초대 편지 발송 시 인센티브 금액이 함께 발송되므로, 1A 그룹은 이메일 받은 후 인센티브를 받게 되고, 1B 그룹은 이메일 받기 전에 이미 인센티브를 받은 상태가 된다. 세 번의 재촉 이메일, 편지를 받게 된다. 이메일은 1주 간격, 메일은 2주 간격이었다.

결과를 보면 인센티브를 먼저 받고 이메일 받은 집단(1B)의 성공률이 가장 높았다(표 6-1 참조). 메일만 받은 집단보다는 메일과 이메일을 함께 받은 집단의 성공률이 더 높았다.

응답자들을 대상으로 설문 편의성에 대한 피드백 조사를 했는데, 응답

자의 다수인 69.5%는 웹서베이를 이용하는 것이 더 좋았다고 응답했다. 양쪽이 차이 없다는 응답이 23.4%, 면접원과 하는 것이 좋았다는 응답이 7%로 나타나 웹서베이가 긍정적인 사용자 경험으로 이끈 것으로 나타났다.

웹서베이가 끔찍한 경험이 되지 않도록 좀 더 설문을 명확히 하고, 웹 화면의 디자인에도 신경을 썼다고 했는데, 응답자들도 서베이의 외관이 괜찮다는 응답이 90%가 넘는 것으로 나타났다.

〈표 6-2〉 HRS의 파일럿 서베이 응답자들에 대한 설문 경험 피드백 조사

Table 5. Respondent Feedback Variables

Question (n)	Percent (n)
Answer feedback questions (439)	87.5% (384)
Experienced technical problems	15.1% (58)
Type of technical problem experienced (of those with 1+) (58)	Other 46.6% (27) Survey froze, had to log back in: 32.8% (19) Survey froze, able to continue: 20.7% (12) Had trouble typing: 8.6% (5) Survey moved slowly: 3.4% (2) Had trouble with prev/next: 3.4% (2) Had trouble selecting answers: 3.4% (2) Had trouble with drop downs: 1.7% (1)
Found the survey easy to complete (383)	Strongly Disagree: 2.9% (11) Disagree: 2.6% (10) Agree: 42.0% (161) Strongly Agree: 52.5% (201)
The help provided was clear and useful (137)	Strongly Disagree: 1.5% (2) Disagree: 5.1% (7) Agree: 51.8% (71) Strongly Agree: 41.6% (57)
The 'look' of the survey was pleasant (372)	Strongly Disagree: 2.2% (8) Disagree: 1.3% (5) Agree: 53.0% (197) Strongly Agree: 43.6% (162)
Self vs interviewer comparison (384)	Better to do it myself: 69.5% (267) About the same: 23.4% (90) Better with an interviewer: 7.0% (27)

자료 : Ofstedal 등(2016).

3. 영국 Understanding Society

영국의 대규모 가구패널조사인 “Understanding Society”는 혁신 패널(innovation panel) 조사라는 이름의 별도 조사를 만들어 도입을 검토 중인 조사방법상의 여러 과제들을 실험연구를 통해 검증하고 있다. 이는 다른 패널조사에서는 보기 힘든 특징이다. “Understanding Society”에서는 첫 조사가 이루어지기 1년 전부터 1,500가구를 잉글랜드, 웨일스, 스코틀랜드에서 표집해서 혁신패널로 조사를 시작했다. 보통 이 패널을 이용해 다른 조사방법이 응답률에 미치는 영향, 다른 질문 방법이 응답 품질에 미치는 영향, 무응답 가구 특성에 대한 조사 등을 실험평가 분석하고 있으며, 지금은 아니지만, 조사 초창기에는 혁신패널이 메인 데이터의 일부로 분석될 수도 있었다고 한다.

웹서베이 관련 실험이 시작된 것은 5차 혁신패널부터이다. 이 조사에서 웹서베이를 검토한 주요 이유는 비용절감이었으며, 응답률 손실 등 위험을 최소화하면서 비용절감을 달성할 수 있을지 확인하기 위한 목적에서 실험을 실시하였다. 즉 웹서베이 실험을 통해 얼마나 많은 사람이 웹서베이를 선택할까? 가구설문을 완성하는 가구는 얼마나 될까? 모든 가구원이 개인설문을 완성하는 가구는 얼마나 될까? 등에 대한 답을 얻고자 하였다.

웹서베이를 기획하면서 어떤 패널가구들이 온라인으로 응답할지에 대해 검토하고자 하였다. Link & Mokdad(2006)의 연구에 따르면, 미국에서는 협조적이었던 가구들이 온라인으로 응답할 가능성이 높으며, 청년, 단독가구, 전문가 계층 등이 좀 더 인터넷 접촉이 빈번하고, 이들은 만나기 어렵다는 점도 있으므로 이들이 정말 웹서베이에 응한다면 응답률을 높이는 측면, 응답률이 오르지 못한다 하더라도 접촉을 위해 많은 방문을 할 필요가 없어 효율성이 높아진다는 장점이 생기기 때문에 이에 대한 검토가 필요했다.

혁신패널을 통해 5차 이후에도 웹서베이 실험이 지속되고 있는데, 웹서베이를 포함하는 혼합조사 방식을 경험한 응답자들의 응답률 장기효과는 무엇일까를 평가하고 있다. 이를테면, 이번에 한 웹서베이 경험이 다음

차수에 덜 참여하게 만들까 더 참여하게 만들까에 대한 확인을 하려면. 다년간 추적이 필요하기 때문이다. 8차 혁신패널까지 웹서베이 실험이 지속되는 중에 있다.

혁신패널을 통해 웹서베이 진행이 조사품질에 영향을 미칠 것인지도 평가되고 있다. 조사도구 변경효과(mode effect)가 존재하는지에 대한 평가이다. 이를테면, 간단한 사실 확인 정도는 별 차이 없을 것이나 면접원 도움이 필요할 수 있는 산업이나 직업에 대한 상세한 묘사 같은 질문에 대한 응답은 차이가 있을 수 있다. 또한 민감한 정보도 그럴 가능성이 있다. 웹의 비대면 특성상 사회적으로 민감한 질문에 좀 더 솔직한 대답을 할 가능성 있다. 이와 같은 전반적인 점검을 위해 혁신패널을 통한 평가연구가 필요하였다.

다만, 전체 조사 대상 가구원을 모두 웹서베이로 성공하지 않는 한, 실제 돈을 절약하는 효과는 별로 없을 것이라 판단하고 있었으며, 웹서베이가 주 모드가 되면 생물학적 정보수집, 행정정보 연결을 위한 동의 확보 같이 개인 대면접촉에서 더 쉬운 것들을 장기적으로 못하게 될 가능성도 고려해야 한다는 점을 알고 있었다. 이 때문에 조사를 위한 하나의 도구로 웹을 활용하는 방법도 생각하고 있었다. 특히 거절 가구, 개인에 대해서만 웹서베이를 활용하는 방안이 그런 사례이다. 다른 방식으로 조사하는 것을 거절한 경우 그럼 웹으로라도 해주겠는지를 타진하는 것도 검토하였다.

구체적으로 5차 혁신패널의 혼합조사 설계를 보면, CAPI만 하는 표본과 웹서베이+CAPI로 진행하는 표본으로 집단을 나누었고, 전체 실험 연구 표본의 1/3은 면접조사, 나머지는 웹서베이 + CAPI 혼합모드로 진행하였다. 혼합조사(웹서베이 + CAPI)에는 1,022가구가 할당되었다.

혁신패널의 대면면접조사 시작 13일 전 웹서베이에 응할 것을 요청하였으며, 대면면접조사 시작 이전에 완성되지 않은 모든 조사대상가구는 대면면접조사로 변경되도록 하였다. 다만, 온라인으로 완성하겠다는 의사를 밝힌 가구는 그대로 완성하게 놔두었다고 한다.

웹서베이에 응하게 하기 위해 인센티브를 포함한 초대편지를 발송하였다. 편지에는 URL과 ID가 포함되었으며, 편지는 이메일 주소를 가진 모

든 사람에게는 이메일로도 추가 발송되었다. 이전 차수 조사에서 개인 목적으로 인터넷을 일상에서 사용하지 않는다고 응답한 사람들에게는 면접조사 기회가 있다는 점을 편지에서 언급했다.

응할 것을 요청하는 단계에서도 실험을 하였다. 이 실험을 통해 만약 웹서베이가 실제 실시될 경우 얼마나 많은 가구가 응할지를 확인할 수 있고, 이 중 완성에 실패할 가구는 얼마나 될 것인지와 같은 정보를 얻을 수 있기 때문이다. 요청 이메일도 금요일날 반, 월요일날 반을 나눠서 보내는 실험도 실시하였다. 이를테면 주말 작성을 편하게 생각하는 사람들이 많은지를 발송 요일에 차별성을 두어 응답률 차이가 있는지 확인하려는 시도였다.

세 번의 이메일 리마인더가 3일 간격으로 발송되었다. 2주 후에도 웹서베이를 완성하지 않은 경우 우편으로 리마인더가 보내지고 면접원이 CAPI 인터뷰를 수행하기 위해 방문하기 시작한다. 다만, 이 시점에서도 조사 완료 시까지 응답자들이 원한다면 CAWI를 할 수 있다.

웹으로 설문을 전혀 못 한 가구부터 부분완성 가구까지 전부 대면면접으로 전환하였다. 그러나 원하는 가구는 웹으로 할 수 있도록 체크하면서, 그렇지 않은 가구부터 면접조사를 실시하였다.

성인 자기기입방식도 대면면접가구에 대해 실시하였다. 대면면접가구 모두를 임의표집방식으로 절반으로 나누어 컴퓨터를 이용한 자기기입식(Computer Assisted Self Interview : CASI)과 자기기입식 종이설문을 각각 받도록 하였다. 원래부터도 Understanding Society는 특별히 민감하다고 생각되는 설문에 한해 자기기입식 설문을 제공했었다.

각 가구의 한 명은 가구 설문 응답을 요청받으며, 16세 이상 모든 가구원은 개인 면접을 요청받도록 되어 있다. 자기기입식 질문지도 완성하도록 요청받는다. 첫 번째로 로그인한 가구원은 가구원 변동 설문지를 응답하도록 요청받으며, 누가 소득, 소비를 책임지는지 식별하도록 요청받는다. 가구 설문은 소득, 소비를 책임지는 사람들에 의해 완성된다. 이를 책임지는 사람들은 웹서베이에 들어오면 가구 설문부터 완성한 후 개인 설문으로 넘어간다.

청소년 조사는 부모가 먼저 조사된 다음, 자녀에게도 설문에 응할 것을

권유하는 초대 이메일을 부모에게 보내는 것으로 시작한다. 이 이메일을 자녀에게 포워드하면 자녀조사가 시작된다(10~15세 청소년을 위한 youth questionnaire가 별도로 있음).

웹설문지는 CAPI 설문에 기반하는데, 면접원이 없는 만큼 설문의 원활한 진행을 도와주는 showcard 같은 것은 없다. 대신, 헬프 메뉴를 만들어 질문에 대한 모호점을 해소할 수 있게 하였다.

웹서베이는 스마트폰 같은 작은 모바일 디바이스에는 적합하지 않아, 이런 것으로 접근할 경우 컴퓨터를 이용해 설문에 응답해 달라는 페이지로 이동하게 된다.

어떻게 하면 웹 응답률을 높일까를 테스트하기 위해 추가 실험을 했다. 가구의 절반은 추가적인 인센티브를 받는 실험이다. 즉 첫 초대편지를 받고 2주 안에 전 가구원이 웹서베이를 완성할 경우 그들 각각은 추가로 5파운드를 받는다. 이 사실은 이 할당에 속한 모든 가구에 보내는 초대편지에 언급되어 있다.

〈표 6-3〉 5차 혁신패널 평가 결과 1

Proportion of households (*n*)	Original sample			Refreshment sample
	Total sample	Wave 4 respondent households	Wave 4 non-respondent households	Total sample
Total sample	18.5 (*728*)	21.0 (*618*)	4.5 (*110*)	34.3 (*315*)
£5	17.5 (*389*)	19.5 (*330*)	6.8 (*59*)	—
£10	21.2 (*311*)	24.7 (*263*)	2.1 (*48*)	23.0 (*87*)
£20	—	—	—	37.1 (*105*)
£30	—	—	—	42.9 (*111*)
P	0.21	0.11	0.22	0.00

주 : The analysis by value of the unconditional incentive excludes *n*=40 split-off households. *P*-values from Wald tests of the equality of mean completion rates between incentive groups, adjusted for sample design.

자료 : Jäckle, Annette, Peter Lynn, and Jon Burton(2013), "Going Online with a Face-to-Face Household Panel: Initial Results from an Experiment on the Understanding Society Innovation Panel", Institute for Social and Economic Research, University of Essex.

〈표 6-4〉 5차 혁신패널 평가 결과 2

Table 4: Household Response Rates

	Original sample									Refreshment sample		
	Total			IP4 responding			IP4 non-responding					
	F2F	MM	*P*	F2F	MM	*P*	F2F	MM	*P*	F2F	MM	*P*
HH response rate	78.0	74.3	*0.22*	84.1	81.1	*0.29*	32.6	36.4	*0.66*	85.1	81.9	*0.45*
Complete HHs	58.0	50.3	*0.02*	63.9	55.7	*0.02*	14.0	20.0	*0.43*	60.1	66.0	*0.26*
Partial HHs	20.1	24.0	*0.13*	20.2	25.4	*0.07*	18.6	16.4	*0.72*	25.0	15.9	*0.01*
Non-contact	6.9	5.8	*0.49*	5.6	2.9	*0.04*	16.3	21.8	*0.45*	5.4	6.7	*0.63*
Refusal	13.2	17.4	*0.13*	8.7	14.1	*0.05*	46.5	36.4	*0.22*	8.3	9.5	*0.68*
Other non-response	1.9	2.5	*0.59*	1.6	1.9	*0.68*	4.7	5.5	*0.84*	1.2	1.9	*0.56*
N	364	728		321	618		43	110		168	315	

주 : F2F = face-to-face; MM = mixed modes; HH = household. *P* = *P*-values from Chi^2 tests adjusted for sample design, These notes apply to tables 5 and 6.

자료 : Jäckle, Annette, Peter Lynn, and Jon Burton(2013), "Going Online with a Face-to-Face Household Panel: Initial Results from an Experiment on the Understanding Society Innovation Panel", Institute for Social and Economic Research, University of Essex

응답이 얼마나 타당한지 검증하기 위해 웹으로 가구원 변동 설문(household grid)에 응답한 가구 중 200가구는 CATI로 가구원 변동 설문을 다시 조사하였다. 이를 조사한 전화 면접원은 웹으로 응답된 정보 없이 재조사하도록 하여 사전 응답에 치우치는 편향을 통제하였다.

인센티브 금액이 높으면 웹으로 응답할 가능성이 증가한다. 가구 응답률을 보면 혼합조사 방식(표의 MM)이 통계적으로 유의한 더 낮은 가구 완전 응답률을 보였다. 완전 응답은 전체를 다 응답한 경우이다. 일부 표본의 경우 혼합조사 방식에서 더 높은 거절률도 나타났다.

조사결과를 보면 개인 응답률도 유의하게 낮게 나타났다. 전체적으로 응답률에 대한 결론을 정리하면 혼합 조사 방식이 더 낮은 응답률을 야기했다는 것이다. "ip4 responding"은 4차 혁신패널에 응답한, 즉 지난 차수 조사에 응답한 가구(개인)를 의미한다. 이전 조사 무응답자의 경우 차이는 무의미했다.

〈표 6-5〉 5차 혁신패널 평가 결과 3

Table 5: Individual Response Rates (including eligible adults in non-respondent households)

	Original sample									Refreshment sample		
	Total			IP4 responding			IP4 non-responding					
	F2F	MM	*P*	F2F	MM	*P*	F2F	MM	*P*	F2F	MM	*P*
Full interview	64.7	58.5	*0.05*	82.8	75.8	*0.04*	18.5	24.1	*0.16*	69.8	71.7	*0.64*
Proxy interview	7.3	4.4	*0.00*	2.6	2.0	*0.50*	20.2	9.4	*0.00*	5.5	1.9	*0.00*
Partial interview	0.1	1.0	*0.02*	0.0	1.2	*0.01*	0.0	0.5	*0.34*	0.0	1.3	*0.03*
Full, proxy or partial	72.1	63.9	*0.01*	85.5	79.0	*0.06*	38.8	34.0	*0.30*	75.3	74.9	*0.93*
Non-contact	7.1	5.8	*0.40*	3.4	2.9	*0.65*	18.0	13.1	*0.20*	5.7	7.1	*0.60*
Refusal	17.2	22.9	*0.06*	8.5	14.2	*0.07*	41.0	43.6	*0.58*	12.9	11.1	*0.59*
Other non-response	3.6	7.4	*0.01*	2.6	3.9	*0.38*	2.2	9.4	*0.00*	6.0	6.9	*0.63*
N	703	1439		495	934		178	406		348	594	

자료 : Jäckle, Annette, Peter Lynn, and Jon Burton(2013), "Going Online with a Face-to-Face Household Panel: Initial Results from an Experiment on the Understanding Society Innovation Panel", Institute for Social and Economic Research, University of Essex

면접원 평균 방문횟수는 2.9회 대 3.7회로 혼합조사에서 낮게 나타났다. 또한, 가구에 인터넷 사용자가 있어야 가구의 웹 응답성향이 상승하는 것으로 분석되었다. 이 페이퍼의 분석에 따르면 21~30세의 혼합조사 응답률도 다른 연령대에 비해 낮은 것으로 분석되었다. 남성, 아이가 있는 가구, 시골지역도 낮을 것으로 분석되었다.

월요일보다 금요일에 웹 초대를 받은 경우, 웹 응답률이 높게 추정되었지만, 통계적으로 유의하지는 않았다.

대략 5가구 중 하나 정도(18.5%)가 웹으로 전적으로 응답했다는 점으로 볼 때 비용절감 가능성 있음을 확인하였다. 웹서베이에 응할 경우 추가 인센티브를 지급한 가구에서 웹서베이 응답률이 22.8%(추가 인센티브 지급) 대 15.5%(추가 인센티브 없음)로 통계적으로 유의하게 높게 나타났다. 항목 무응답률도 면접조사 대 혼합조사가 23% 대 32%로 통계적으로 유의하게 혼합조사가 높게 나타났다. 데이터 품질에 악영향을 미칠 가능성을 확인하였다. 총임금항목 무응답률도 높게 나타났다.

〈표 6-6〉 5차 혁신패널 평가 결과 4

Table 10: Item Non-Response Rates by Sample and Treatment

	Original sample			Refreshment sample		
	F2F	MM	*P*	F2F	MM	*P*
Mean Item non-response (%)	0.23	0.32	*0.01*	0.19	0.27	*0.02*
N	459	859		243	437	
Gross pay Item non-response (%)	10.0	17.7	*0.02*	8.0	18.0	*0.03*
N	229	412		112	206	

주 : *P* : *P*-value from Wald tests of means adjusted for sample design.
자료 : Jäckle, Annette, Peter Lynn, and Jon Burton(2013), "Going Online with a Face-to-Face Household Panel: Initial Results from an Experiment on the Understanding Society Innovation Panel", Institute for Social and Economic Research, University of Essex

5차 혁신패널의 비용절감 가능성은 확인했으나, 응답률과 조사 데이터 품질에 나쁜 영향을 미칠 가능성이 높다는 결과 때문에 6차 혁신패널(2013년 조사)에서도 다시 웹서베이 실험을 수행하였다. 혁신패널은 1차 혁신패널가구원을 추적하며, 원가구에 있든 없든 추적한다. 이들이 새로운 가구를 형성하면 그 회차에는 그 가구원들이 조사에 포함된다. 혁신패널 6차 조사는 1,397가구, 16세 이상 2,767명의 개인이 조사되었다.

웹서베이가 중간에 중단되면 다시 시작할 때 그 자리부터 다시 시작하게 프로그래밍되었다. 조사는 아래 3단계 방식으로 실시되었다.

5차와 동일하게 1/3은 CAPI, 2/3는 혼합조사로 수행되었으며, 임의표집하였다. 5차와 달리 6차에서는 조사기간 종료 즈음에서 혼합모드 그룹에는 전화를 해서 웹서베이를 권한 후 전화면접(CATI)도 가능함을 고지하였다. 전화번호가 없는 가구는 이와 같은 절차를 수행하지 않았다. CAPI 그룹에도 서베이 종료 시점에서 웹서베이에 응하겠냐는 편지를 보냈다. 그러고 나서 이틀 뒤 전화 면접원이 웹서베이나 전화면접에 응해달라는 전화를 했다.

인센티브 실험도 했다. 5차분석 결과는 웹서베이에 더 많은 인센티브를 줄 경우 웹서베이 가능성이 올라감을 보여주었다. 이를 감안해 혼합조사 그룹을 세 그룹으로 나누어 조사 대상 성인당 10파운드, 조사대상 성인에

〔그림 6-2〕 6차 혁신패널의 웹서베이 실험 조사 단계 구성

Table 2.1 Data collection timetable	
	Timing
Phase 1 – CAWI only – Web sample only	2 weeks
Phase 2 – CAPI for unproductive Web sample and all F2F sample. CAWI available alongside for Web sample only	10 weeks + 2 weeks mop-up 7 weeks for CAWI
Phase 3 – CATI mop-up for Web and F2F samples, CAWI available for both samples	4 weeks

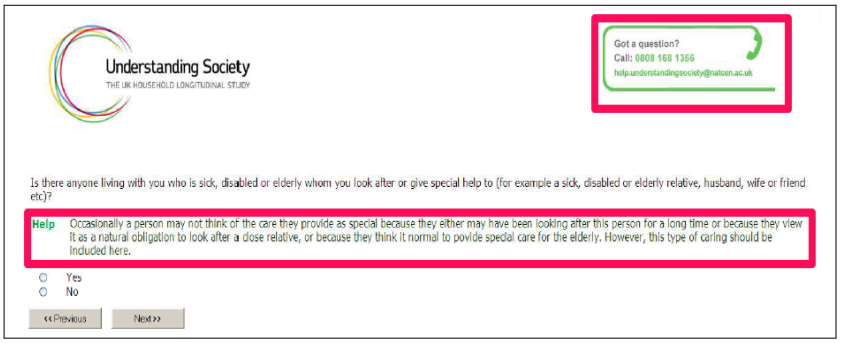
자료 : Agur, Maya, Curtis Jessop, Sarah Kunz, Violetta Parutis, and Martin Wood(2014), "Understanding Society Innovation Panel Wave 6, Technical Report", The Institute for Social and Economic Research, University of Essex.

게 10파운드와 2주 안에 웹으로 전체 가구를 완성할 경우 20파운드 추가 지급, 모든 성인에게 30파운드를 지급하는 세 가지 실험을 실시하였다. 항목 무응답 실험도 했다. 그룹1은 모름 또는 거절 선택지를 그대로 노출하는 그룹(5차 혁신패널 디폴트)이며, 그룹2는 모름 또는 거절 선택지를 선택하면 응답이 중요하니 응답을 해달라는 메시지를 띄우는 그룹이다. 그룹3은 설문의 끝에서 여섯 핵심 질문 중 하나 이상 항목에 무응답이 있는 경우 스크린이 뜨면서 응답하지 않은 핵심질문들에 이 질문들이 매우 중요함을 설명하면서 대답해 주기를 요청하는 그룹이다. 이렇게 세 그룹으로 구분해 어떤 집단의 무응답이 가장 낮은지 실험하였다.

다만, 조사 중에 자녀 설문이 응답률이 낮아서 방식을 바꾸어야 했다고 한다. 자녀 있는 가구에 종이설문을 보내고 이를 자녀에게 완성할 것을 요청함과 아울러 반송봉투까지 넣어서 우편으로 작성된 설문지를 보내줄 것을 요청하였다.

[그림 6-3]은 웹서베이 조사 화면이다. 질문이 있으면 전화하거나 이메일을 보내도록 유도하였고, "헬프"를 노출해 설문의 의미에 대한 설명도 실시하였다.

〔그림 6-3〕 6차 혁신패널의 웹서베이 조사 화면



자료 : Agur, Maya, Curtis Jessop, Sarah Kunz, Violetta Parutis, and Martin Wood (2014), "Understanding Society Innovation Panel Wave 6, Technical Report", The Institute for Social and Economic Research, University of Essex.

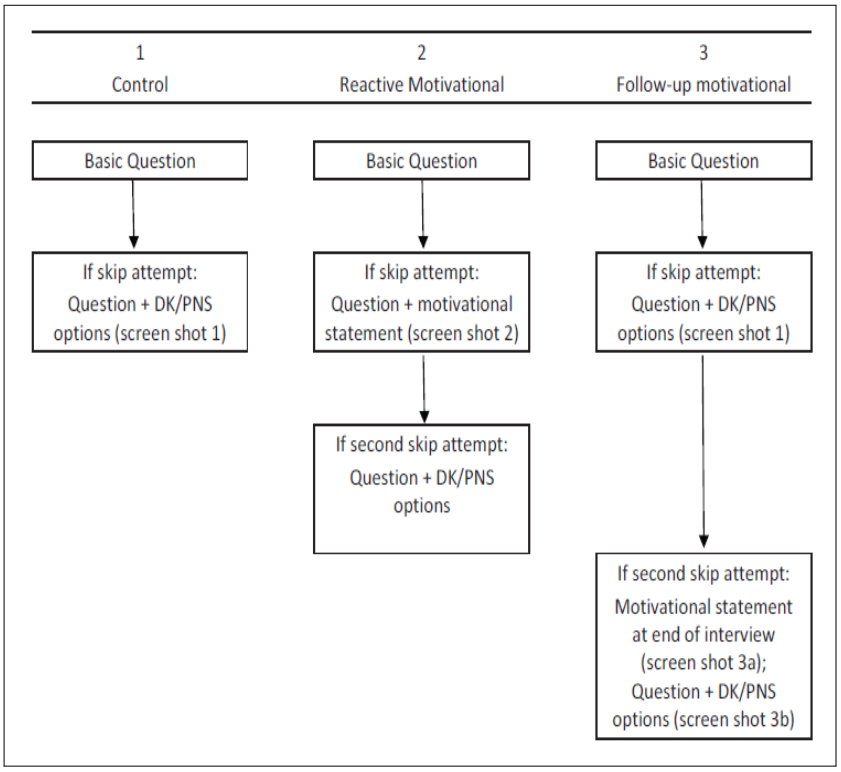
분석결과를 보면, 5차 혁신패널과 달리 6차 혁신패널에서는 결과가 달라졌다. 혼합조사 방식과 CAPI only 사이에 응답률 차이가 각각 81.9%대 82.1%로 없어졌고, 완전한 가구응답 측면에선 오히려 혼합조사가 더 높은 응답률을 보였다(혼합조사 65% 성공, 대면면접조사 61.5% 완성). 개인 응답 측면에서도 혼합조사가 응답률 73%로 대면면접조사 70.7%보다 높게 나타났다.

이 조사에서는 항목 무응답과 관련된 실험도 있었다. 다시 구조를 설명하면, 다음과 같은 구조이다. 1=응답 안 하고 next를 눌러 스킵하려고 하면 그제서야 다시 질문하면서 모름/무응답이 있는 선택지 뜨는 방식이며, 2=응답 안 하고 next를 눌러 스킵하려고 하면 제대로 된 응답이 중요하므로 꼭 응답을 부탁드린다는 내용의 권유문장이 뜨는 방식이다. 3=일단 스킵하게 하고 설문이 종료될 때 응답 안 된 설문들을 보여주면서 제대로 된 응답을 해줄 것을 권유하는 방식이다.

실험 결과는 <표 6-7>에 소개되어 있는데, 2번 방식이 효과적인 것으로 나타났다. 대면면접과 유사한 수준의 무응답률을 보이는 것으로 나타났다.

이와 같은 Understanding Society의 실험은 8차 혁신패널까지 지속되었다. 8차 혁신패널은 실험 할당에서 이전과 중요한 차이가 있다. 혼합조사 응답 가능성이 가장 낮은 것으로 분석된 하위 20% 표본가구는 CAPI

〔그림 6-4〕 6차 혁신패널의 웹서베이 항목 무응답 실험 설계



자료 : Al Baghal, Tarek, Peter Lynn(2015), "Using Motivational Statements in Web-Instrument Design to Reduce Item-Missing Rates in a Mixed-Mode Context", *Public Opinion Quarterly* 79(2), pp.568~579.

가구로 할당하며, 나머지 80% 가구를 반반씩 혼합조사 가구와 CAPI 가구로 할당하였다. 따라서 전체 가구 중 60%는 CAPI로 조사되며 나머지 40%만 혼합조사로 할당된다. 응답가능성이 낮은 가구 여부는 5~7차에 조사된 결과에 의거 결정되었다.

3~6월에는 CAWI만 시도하였다. 임의표집된 100가구는 좀 더 빨리 3월 초부터 조사해 3월 중순까지 조사를 끝냈다. 기술적으로 발생할 수 있는 문제들을 점검하는 것이 목적이었다. CAWI 메인조사는 3월 중순부터 시작되었다. 6~9월 중순은 CAWI 실패 또는 불완전하게 조사된 가구와 CAPI 할당가구에게 CAPI로 응할 것을 요청하고 조사하는 기간이었다.

조사 관련 메일은 5월 말 발송되었다. CAWI로 조사될 수도 있는데, 6월 말부터는 무조건 CAPI로만 조사되었다.

〈표 6-7〉 6차 혁신패널 웹서베이 항목 무응답 실험 평가 결과

Table 2: Weighted item nonresponse rates by treatment; mixed mode and face-to-face designs

	Mixed Mode Outcomes			
	Control	Reactive	Follow-Up	Face-to-Face
Gross pay	0.150	0.129	0.185+	0.118
Interest/Dividends	0.209+	0.146	0.192+	0.136
Net Profit	0.583+	0.452	0.228-	0.404
Hourly Pay	0.108	0.057	0.166+	0.061
Marital Status Change	0.016	0	0	0
Total	0.180+	0.127	0.156+	0.120

주 : + indicates mixed mode outcome significantly greater than face-to-face only outcomes at $p<0.05$
- indicates mixed mode outcome significantly less than face-to-face only outcomes at $p<0.05$

자료 : Al Baghal, Tarek, Peter Lynn(2015), "Using Motivational Statements in Web-Instrument Design to Reduce Item-Missing Rates in a Mixed-Mode Context", *Public Opinion Quarterly* 79(2) pp.568~579.

〔그림 6-5〕 8차 혁신패널의 웹서베이 실험 조사 단계 구성

Table 2.1: Phases of fieldwork design

	Phase 1: CAWI only	**Phase 2: CAPI interviewing**	**Phase 3: CAWI/CATI mop-up**
	May 2015	June –September 2015	Late September 2015
CAWI-first households	Invited to complete via **CAWI**	Incomplete cases invited to complete via **CAPI**	Incomplete cases invited to complete via **CAWI** or **CATI**
CAPI-first households	-	Invited to complete via **CAPI**	Incomplete cases invited to complete via **CAWI** or **CATI**

자료 : Hanson, Tim, Peter Matthews, and Alice McGee(2016), "Understanding Society Innovation Panel Wave 8, Technical Report."

9월 중순에서 말까지는 둘 다 실패, 또는 불완전하게 조사된 가구를 위해 CAWI 또는 CATI에 응해 줄 것을 청하고, 본인들이 원하는 방식으로 조사되게 해주는 기간으로 설정되었다. 2주간 조사하였다.

여기에서도 유인 실험이 실시되었다. CAWI 그룹은 조건 없이 10파운드를 받는 그룹, 조건 없이 30파운드를 받는 그룹, 10파운드+모든 조사 대상 성인이 CAWI 조사 기간(3~6월 말) 동안 CAWI로 응답할 경우 20파운드 추가하는 세 그룹으로 균등 분할되었다. 분석결과를 보면, 응답률이 가장 높은 그룹은 무조건 30파운드를 주는 그룹이었다. 89.2%의 응답률로 나타났다. 20파운드 조건부가 붙는 그룹은 82.1%의 응답률을 보였으며, 가장 응답률이 낮은 그룹은 10파운드 받는 그룹으로 78.5%의 응답률을 기록했다.

이와 같은 결과는 혼합조사 방식이라는 특성의 반영으로 보인다. 전체를 CAWI로 하지 않아도 CAPI나 CATI에 응하기만 하면 30파운드를 받

〔그림 6-6〕 8차 혁신패널의 웹서베이 유인 실험 설계 결과

Table 8.4: Household response by incentive group

Base: CAWI-first households	**£10 incentive**	**£10 + £20 on full household completion**	**£30 incentive**	**Total**
Any productive	78.5% 183	82.1% 202	89.2% 223	83.4% 608
Fully productive	55.4% 129	66.3% 163	72.8% 182	65.0% 474
Partially productive	23.2% 54	15.9% 39	16.4% 41	18.4% 134
Any unproductive	21.5% 50	17.9% 44	10.8% 27	16.6% 121
HH Grid or HH Qnr only	2.1% 5	2.0% 5	2.4% 6	2.2% 16
Refusal	8.2% 19	6.1% 15	2.8% 7	5.6% 41
Non-contact	6.4% 15	6.1% 15	4.0% 10	5.5% 40
Other unproductive	4.7% 11	3.7% 9	1.6% 4	3.3% 24
Bases	*233*	*246*	*250*	*729*

자료 : Hanson, Tim, Peter Matthews, Alice McGee(2016), "Understanding Society Innovation Panel Wave 8, Technical Report."

을 수 있으니까 그렇다. 그러나 전 가구원 CAWI 조건부로 30파운드를 받을 수 있는 집단은 CAWI가 아니라면 10파운드밖에 받을 수 없는 구조니까 낮았을 것이다.

다만, CAWI로 실제 전체 설문을 완성한 가구 비중은 조건부 30파운드 집단이나(10+20) 무조건부 30파운드 집단에서 큰 차이가 없는 것으로 나타났다(각각 성공률 48.4% 대 50%).

응답률을 끌어올리는 그 자체가 목표라면 가구에 지급하는 수당을 무조건부로 인상하는 것이 가장 좋겠지만, 그렇게 하기엔 재정 여력이 부족한 상황이라면, 여기에 더해 비용절감이 목표라면 CAWI로 설문을 완성시키는 것이 중요하므로 전 가구원 조건부 인센티브체계를 만드는 것도 대안이 될 수 있음을 시사하는 결과이다.

4. 기 타

가. The Netherlands Kinship Panel Study

이 조사는 18~79세를 대상으로 하는 60분 정도 길이 되는 긴 설문을 가진 패널조사이다. 3차 조사에서 웹서베이를 도입했다. 1, 2차 조사는 CAPI였다, 3차 조사에서는 우선 웹서베이로 조사해 달라고 하였으며, 약 40% 정도가 이에 응했다고 한다. 응하지 않으면 CATI(컴퓨터 기반 전화면접)로 넘어가는데 20% 정도가 응했다고 한다. 최종적으로 개인 인터뷰가 CAPI 방식으로 진행되었는데 약 13% 정도가 이 방식으로 조사되어 2차 조사 응답자 중 73%의 응답률을 거뒀다고 한다. 이와 같이 순차적으로 다음 조사방식으로 넘어가는 방식의 복수조사도구 조사설계를 sequential mixed-mode design이라고 부른다(Dykstra, 2011). 이 방식의 결과, Per Liefbroer(2011)은 응답자당 비용이 2차 조사 195유로에서 3차 조사 115유로로 떨어졌다고 보고하였다.

다만, 이 조사는 가구당 한 명만 응답하면 되는 조사라는 점이 중요하다. 이 때문에 성공적이었을 수 있다. 노동패널이나 BHPS 같은 가구패널조사는 보통 전체 가구원이 응답해야 한다.

스위스 가구패널조사는 2010년 이래 거절 가구에 한해 CAPI 또는 CAWI를 실시했다(Voorpostel and Ryser, 2011). CATI가 실패한 1,962패널가구에게 웹서베이를 부탁한 결과 87가구가 웹을 선택했다. 그러나 실제 완성한 것은 단지 43가구에 불과했다. “The Netherlands Kinship Panel Study”도 2차 조사 시 실패가구에 대해 웹서베이를 제안해 1.6%를 조사 성공했다고 한다(Dykstra, 2007). 만약 만나야 하는 일부 가구원이라도 웹에 응해 준다면 추가방문의 필요성이 낮아지니 이런 관점에서 시도해 볼 수 있겠다.

나. 독일 SOEP

CAPI, PAPI 같은 면접방식을 주 모드로 하되, 수행과정에서 응답자 선택 범위 확장 차원에서 2014년부터 본조사에 CAWI를 실시하였다. 이 조사에서는 원래부터 조사로 여러 모드를 쓰고, 특히 PAPI(종이설문을 위한 대면면접조사)를 함께 사용하기 때문에 과거 정보를 연동하는 방식의 조사는 CAPI에서도 수행하지 않는 특색이 있다. 따라서 과거 정보연동을 CAWI에서 어떻게 할 것인지와 같은 이슈는 없었다.

그런데, CAWI는 2014년부터 새롭게 표본으로 만든 확장 패널에서만 실시하였다. 이 표본에서 웹서베이 도입을 위한 준실험 연구가 수행되었다. 새로운 표본을 통합하는 과정에서 예산절감의 필요 때문에 CAWI 실험이 실시되었다.

이 표본은 FiD, 독일의 가족이라는 이름으로 이미 2010년부터 독립적으로 조사되는 것이다. SOEP과 비교가능하게 발전하고 있었다. 기존에는 CAPI로 조사되었다. CAWI-CAPI 순서형 조사설계를 한 결과, 가구당 비용이 대략 30%쯤 감소한 것으로 나타났다.

제6절 웹서베이 도입을 위한 파일럿 서베이 제안

선행 연구 결과로 볼 때 웹서베이 도입 시 비용절감은 장기적인 목표는 될 수 있어도 향후 몇 년 내의 목표는 되기 어려울 것으로 보인다. 비용절감이 가능하려면 면접원의 방문 없이 웹서베이만으로 성공하는 가구가 많아야 하는데, 가구와 개인 모두를 조사해야 하는 노동패널의 특성상 넘어야 할 관문이 많을 것이기 때문이다.

다만, 특정 인구집단, 이를테면 청년이나 1인 가구처럼 대면 자체가 어려운 집단에 대해서는 웹서베이를 병행하는 것이 응답률을 높이는 데에 도움이 될 수 있다. 또한, 앞서 본 사례에서 나오는 것처럼 응답거절 가구에 대해 웹서베이에는 응할 수 있는지 조사 권유하는 수단으로 웹서베이를 활용하는 것은 응답률 개선이라는 측면에서 유용할 수 있다.

디지털화가 나날이 진전되는 상황에서 조사 비용 절감을 고려하지 않더라도 디지털화된 수단을 통한 조사는 노동패널이 추구해야 할 방향으로 보인다. 다른 대규모 가구패널조사에서 실험적으로든 부분적으로든 도입하는 것도 조사비용 고려뿐 아니라 디지털화가 진전되는 조사 경향과 인구집단에 따라서는 웹을 이용한 방식에 더 잘 반응할 수 있는 측면을 모두 고려하기 때문이다. 따라서 노동패널도 웹서베이에 대한 현실성 검토를 하는 것은 조사의 발전을 위해 꼭 필요한 것으로 보인다.

현재 단계에서 노동패널의 웹서베이는 기존 조사도구를 대체하기보다는 부분적으로 보완하는 수준에서 도입을 검토하는 것이 현실적인 것으로 보인다. 그렇다면, 도입 검토 단계에서의 웹서베이 파일럿 테스트는 무엇을 목표로 해야 하는가? 선행 연구를 종합해 보면, 웹서베이가 어떤 인구집단에서 좀 더 응답 개선을 끌어내는 데에 효과적인지, 웹서베이로 조사를 했을 때 기존 조사 방법과 조사 결과가 다른지, 즉 조사도구 혼동효과(mode effect)가 있는지를 검토하는 것이 일차적인 목표여야 할 것으로 보인다.

또한, 단기에는 웹서베이를 도입하더라도 조사 실패 가구에 우선적으

로 투입한다면 응답 개선 효과보다는 조사도구 혼동 효과가 존재하는지를 검증하는 것이 좀 더 중요한 목표일 것으로 보인다.

이 절에서의 파일럿 조사 제안은 따라서 일차적으로는 조사도구 혼동 효과가 존재하는지 여부를 검증하는 것을 목표로 하고, 가능하다면 어떤 인구집단에서 좀 더 응답 개선 가능성이 있는지까지 검토하는 것을 목표로 삼을 수 있다면 더욱 좋을 것으로 보인다. 조사도구가 바뀌면서 나타나는 효과보다 설문이 혼자 해결하기엔 어려워서, 즉 단순 이해도 차이로 결과가 달라질 수도 있을 것이다. 따라서 FGI를 통해 이해도를 정성적으로 파악해 보는 것도 병행되어야 할 것으로 보인다.

다만, 영국에서 웹서베이 실험만 네 차례에 걸쳐 꾸준히 하고 있는 것처럼 이번 파일럿 조사 결과로 웹서베이 도입 여부를 결정하는 것은 신중한 결정이 아닐 수 있다. 다년간의 연구를 통해 웹서베이의 장단점을 파악한 후 노동패널 본조사에 도입하는 것이 현명할 것으로 보인다.

파일럿 조사를 노동패널 표본에서 실행하기에는 위험부담이 따른다. 따라서 이번 조사는 노동패널 본표본이 아닌 별도의 표본을 통해 웹서베이가 노동패널 설문구조에서 실제 가능한지, 웹서베이와 CAPI 결과가 다른지만 확인하는 것을 목표로 삼자는 제안을 한다.

1) 비용절감을 목표로 생각한다면 예비테스트를 통해 얼마나 많은 가구가 웹서베이에 응할지, 웹서베이에 응한 가구 중 면접원 동원이 없이 설문을 완전히 완성한 가구가 얼마나 될지 테스트되면 좋겠다. 그러나 노동패널 본표본이 아닌 한 이에 대한 실험은 불가능하다. 파일럿을 위한 표본이 노동패널 표본과는 특성이 다를 수 있고, 1회성 조사의 응답을 연속적으로 조사되는 패널조사의 응답과 비교하기는 어렵기 때문에 파일럿으로 이를 테스트한다 해도 비교가능성이 떨어지기 때문이다.

2) 웹서베이는 혼합조사 방식으로 테스트되어야 한다. 왜냐하면 다른 외국 사례를 참고할 때 웹서베이로 모든 조사가 완료되지 않는 상당수의 가구들이 존재할 것이기 때문이다. 즉 웹서베이 할당 집단은 웹서베이 응답을 거절하거나 웹서베이를 하다가 중단할 경우 CAPI 면접원이 투입된다. 다만, 이는 웹서베이 파일럿 테스트의 목표가 설문의 이해도, 조사혼동효과일 때 그다지 바람직한 방법은 아니다.

3) 특정 인구집단만 실험하는 것이 아니라 전체 가구 및 가구원을 대상으로 노동패널과 유사한 환경에서 테스트를 실시한다. 청년이나 단독가구처럼 만나기 어려운 대상 집단에 한해 테스트할 수도 있지만, 이들에게서 나타나는 응답패턴이 전체 가구와는 다를 수 있어 웹서베이가 조사품질에 미칠 영향을 테스트해 보는 목적이라면 전체를 대상으로 테스트를 진행하는 것이 좋을 것으로 판단된다.

이 정도의 고려를 한다면, 전체 가구, 가구원을 대상으로 하여 파일럿 조사를 하되, 노동패널 본표본이 아닌 별도 표본을 대상으로 하고, 대신 응답률 평가는 포기하고 조사도구 혼동효과와 조사 이해도 수준 확인 정도를 목표로 파일럿 조사를 하도록 제안하는 것이 본 연구의 잠정적 제안이다. 대신 CAPI 가구와 웹서베이 가구는 무작위 할당하여 실험 연구 틀을 통해 평가를 진행한다.

제 7 장
웹을 활용한 조사의 효과에 대한 실증연구[43] : 노동패널 CAWI 실험 데이터의 분석

제1절 서 론

한국노동패널은 2017년 현재 20차 조사를 진행하고 있을 만큼 양적으로 질적으로 많이 성장하였다. 한국노동패널조사는 동일한 가구를 계속적으로 조사하여 한 개인의 일생 동안의 출생부터 사망까지의 교육, 혼인, 일자리, 건강 등의 총체적인 내용을 담고 있다. 따라서 표본의 유지율이 패널 조사의 승패를 좌우한다고 할 수 있다.

한국노동패널 조사 17차 조사 원가구 유지율은 1차년도 응답가구 5,000가구 대비 69.0%로 다른 패널에 비해 성공적으로 유지되고 있다. 또한 패널의 노령화를 막기 위해 매년 조사가구의 35% 이상의 분가가구를 조사하고 있다. 그럼에도 불구하고 비성공 가구들은 해마다 사망, 이민 등의 필연적 사유로 발생할 수밖에 없다. 그 외 비성공 사유를 살펴보면, 노동패널 초기에는 강력거절이 주를 이뤘으나, 최근 5년 동안은 이사추적실패

43) 본장은 노동패널 조사도구 전환(PAPI(Paper Assisted Personal Interviewing)에서 CAPI(Computer Assisted Personal Interviewing)로의 전환)을 위해 실시한 실험연구를 바탕으로 작성되었던 최효미 외(2012), 『패널자료의 품질 개선을 위한 조사기법 연구 - 컴퓨터를 이용한 조사(CAPI) 방법』의 방법론을 이용하여 분석한 것이다.

가 주요 원인으로 꼽히고 있다. 이는 패널이 조사를 거부하는 것이 아니라, 연락처가 없어서 찾아가지 못하는 것이기 때문에, 이를 해결한다면 패널 유지율을 더욱 안정적으로 가져갈 수 있을 것이라 기대된다. 따라서 집주소에 비해 상대적으로 변동이 덜한 휴대전화번호나 이메일을 수집하는 것이 필요해졌다. 또한 과거에 비해 낯선 사람을 대면하는 것을 꺼리는 분위기까지 겹쳐, 특히 1인가구 젊은 연령의 패널의 응답률은 계속 떨어지고 있는 실정이다.

이에 대한 대안 중 하나가 웹서베이가 될 것이다. 그러나 웹서베이의 필요성에도 불구하고, 웹서베이는 면접원의 설명 없이 조사가 이루어지기 때문에, 응답자가 설문을 올바르게 이해하지 못하는 문제가 발생할 수 있다. 또한 본인이 설문에 응했는지, 대리로 응답하였는지를 확인할 수 없는 단점들이 있다.

따라서 노동패널팀에서는 웹서베이 실험연구를 실시하여, 자료의 체계적인 분석을 통해 웹서베이의 도입을 검토하고자 한다. 본 보고서에는 웹서베이를 CAPI(Computer Assisted Personal Interviewing)와 대조되는 용어인 CAWI(Computer Assisted Web Interviewing)로 정의하여 사용하였다.

CAWI 실험의 목적은 ① 면접원이 없어도 응답자가 설문을 잘 이해하고 응답하는가? ② 조사방법에 따라 설문응답의 일관성이 유지되는가? 즉 주요 문항들에서 CAPI와 CAWI의 응답 값이 차이가 나지 않는가를 살펴보는 것이다. ③ 조사도구의 변화에도 현재의 조사 성공률을 유지할 수 있을까? 즉 응답자가 설문을 잘 이해하고, CAPI조사와 CAWI조사 결과의 차이가 존재하지 않더라도, 조사 성공률이 현재 CAPI 수준에 미치지 못한다면, 전면 도입은 어렵다는 것이 전제되어야 한다.

제2절 노동패널 CAWI 실험 설계 개요

노동패널 CAWI 실험은 다음과 같은 순서로 진행되었다.

첫째, 해외패널에서 활용되고있는 CAWI 조사에 대해 자료를 수집하고, 본 실험에서 활용할 수 있는 방안들과 시사점을 살펴보았다.

둘째, 실험설계를 위한 CAWI 프로그램을 개발하였다. 현재 노동패널은 CAPI 조사를 위해 네덜란드 통계청이 개발한 BLAISE 저작권을 해마다 갱신하여 이용하고 있다. 2009년 당시 BLAISE를 처음 사용할 때에는 한국에 컴퓨터를 기반으로 하는 설문조사가 보편화되어 있지 못하였기 때문에 검증된 외국 프로그램을 이용하는 것이 최선의 선택이었다. 그러나 현재는 많은 조사업체가 이미 웹 기반의 자체 프로그램들을 개발하여 사용하는 것이 보편화되었고, 그에 대한 검증은 많은 조사를 통해 이루어졌다. 이에 외국 프로그램인 BLAISE를 고집하기보다는 현재 노동패널 실시업체인 한국리서치에서 새로운 프로그램을 개발하고 이를 예비적으로 시행해 보는 기회로 CAWI 실험을 이용하고자 하였다. 따라서 본 조사는 CAWI의 도입가능성 파악 목적 외에도 내부적으로는 새로운 프로그램이 오류 없이 작동하는지도 검토하고자 하였다. 프로그램 개발기간은 약 4개월 정도가 소요되었다.

셋째, 실험 설계 조사대상 가구를 확정하여 조사를 실시하였다. 웹조사의 도입여부를 판단하기 위해 기존 CAPI조사와 실험 CAWI조사의 응답차이를 비교 분석하여야 했다. 따라서 실험집단 CAWI 300명과 비교집단 CAPI 150명이 속한 가구의 15세 이상 가구원 전수를 목표로 설정하였고, 조사는 2016년 10 ~ 12월 사이에 이루어졌다.

CAWI 전환을 위한 실험 표본 표집 절차는 [그림 7-1]과 같다.

예비 조사를 위해 한국리서치에서 보유하고 있는 MS(Master Sample) 패널을 모집단으로 이용하였다. MS 패널은 한국리서치에서 전국의 국민을 지역, 성, 연령, 직업, 학력, 소득 분포에서 통계적으로 대표할 수 있도록 계획하여 모집한 응답자 패널로 약 34만여 명이 등록되어 있다. MS 패널은 한국리서치 설문조사에 응하겠다고 지원하여 여러 설문을 경험한 사람들이다. 따라서 기본적으로 노동패널 본조사 패널들보다 웹조사에 능숙하기 때문에 응답을 하는 데 있어 오류나 어려움 없이 진행할 수도 있다. 또한 조사에 응하겠다고 미리 신청한 사람들이므로, 일반인보다 조사에 대한 거부감이 적어 조사 응답률이 높을 수 있음에 유념해야 한다.

노동패널팀은 MS 패널 중 무작위로 20,931명의 사람에게 메일을 보내, 노동패널조사에 응답해 줄 응답자를 공모하였다. 이에 4,224명이 사전조사 설문[44]에 참여하였고, 목표 표본의 2배수로 표본을 추출하였다.

추출 원칙은 지역별 가구원수 비율을 한국노동패널 19차 조사 대상가구와 동일하게 할당하는 것이다. 이에 따라 실험 조사에 응답하겠다고 신청한 사람에게 사전에 거주지와 가구원수를 질문하여 할당에 필요한 정보를 습득하였다. 목표집단 CAWI 300가구, 비교집단 CAPI 150가구로, 지역별 가구원수별로 할당 표집하였다.

넷째, 실험조사 결과를 분석하였다. 주요 분석 항목은 기본적인 조사정보(성공률, 응답일수, 응답시간, 독려 횟수, 중단 시점)와 항목응답률, 분포의 차이, 응답 값 차이, 항목무응답률, 설문유형에 따른 차이, 직산업 분류 차이 등이다.

각각의 항목 내에서는 가구의 소득이나 자산, 생활비 등 가구설문에 포함된 변수와 개인의 근로소득이나 일자리 정보와 같은 개인설문에 포함

〔그림 7-1〕 표본 추출 과정

(단위 : 명)

MS 패널	➡	공모이메일 발송자	➡	사전조사 참여자1)	➡	2배수 표본 선정2)			➡	목표 표본3)		
						전체	CAPI	CAWI		전체	비교집단 (CAPI)	실험집단 (CAWI)
344,430		20,931		4,224명		816	280	536		450	150	300

주 : 1) 사전조사 설문을 시작한 사람이 4,224명 중 설문을 완료하지 않거나 지역별 가구원수 할당이 과도 표집된 집단의 참여자 3,408명을 제외함.
2) 816명 중 웹조사 희망자는 601명, 대면조사 희망자는 24명, 조사방법에 상관없다고 응답한 응답자는 191명이었음. 따라서 CAWI는 536명 전체를 웹조사 희망자로 할당하였고, CAPI는 대면조사희망자(24명)와 조사방법에 상관없다고 한 응답자(191명), 그 외 웹조사 희망자 중 65명을 강제 할당하였음.
3) 목표 표본은 조사 응답자가 속한 가구와 해당가구에 속한 15세 이상 가구원 전수임.

44) 사전조사 설문은 거주지, 성별, 가구원수, 가구주와의 관계, 실험 참여 의향, 실험 참여방법(대면면접조사, 웹조사, 둘 다 상관없다), 만 15세 이상 가구원수의 문항으로 구성되어 있다.

된 변수를 함께 분석하였다.

실험 분석에 사용된 자료는 노동패널팀이 한국리서치에 의뢰하여 2016년 10월부터 2016년 12월 사이에 조사된 자료이다.

분석방법은 항목응답률, 항목무응답률과 연속 변수 등을 검정할 때에는 t-검정(t-test)을, 분할표(contingency table)의 경우에는 독립성 검정(x^2검정) 또는 피셔의 정확 검정(Fisher's exact test)을 사용하였다. 또한 분산의 차이가 있는지를 확인할 때에는 분포형태와 관계없이 검정할 수 있는 Levene 검정을 실시하였다.

제3절 CAWI의 효과에 관한 실증분석

1. 조사성공률, 응답시간, 가구원수 대비 성공 가구원수

가. 조사성공률

웹조사를 진행할 경우, 응답률에 어떤 차이가 있는지 파악하기 위해 설문조사 성공률을 분석하였다. 실험조사를 위해 접촉한 가구는 총 816가구(CAPI 280명, CAWI 536명)이다. CAPI는 64%의 조사 시작률을 보이는 반면, CAWI는 92%의 시작률을 보여 CAWI의 조사 시작률이 유의미한 차이로 더 높은 것으로 나타났다.

조사가 시작된 이후 조사완료 성공률을 살펴보면, CAPI는 85%, CAWI는 62%로 유의미하게 CAPI의 조사완료률이 높은 것으로 나타났다. 즉

〈표 7-1〉 조사도구별 접촉가구 대비 조사시작률 차이(가구 설문)

	접촉 가구	조사시작 가구	조사 시작률	S.D	diff[1)]	t-value	Pr>|t|
CAPI	280	178	0.6357	0.4820	-0.2878	-10.99	<0.0001
CAWI	536	495	0.9235	0.2660			

주 : 1) diff는 CAPI 조사시작률에서 CAWI 조사시작률을 뺀 차이를 의미함.

직접적으로 조사를 시작하는 확률은 CAWI가 유의하게 높지만, 최종 조사를 완료하는 확률은 CAPI가 높다는 것을 보여준다.

두 집단 모두 조사에 응하겠다는 신청을 받아서 시작했음에도 불구하고, 조사 시작률에 차이를 보이는 가장 큰 이유는, 조사를 신청한 본인뿐만 아니라, 15세 이상 가구원 모두가 설문에 응해야 한다는 것 때문인 것으로 생각된다. CAPI는 면접원이 방문 전에 전화를 걸어, 사전 약속을 잡는 과정에서 가구원 전체가 모이는 시간을 요청했을 때, 대부분 거절되었다. 그러나 CAWI는 이메일상에 가구원 전원 참여의 공지를 하지만, 일단 설문을 시작하는 경우가 많기 때문인 것으로 생각된다.

가구원 설문은 가구설문보다 조사성공률이 조사도구에 상관없이 모두 높았는데, CAPI 98.8%, CAWI 89.8%로 통계적으로 유의미하게 높았다. 개인설문이 상대적으로 가구설문보다 높게 나온 이유는 이미 가구설문에서 강력거절 가구들이 많이 탈락되었기 때문이라 생각된다.

그러나 접촉가구로부터 조사성공률을 보면, CAWI가 CAPI보다 높기는 하지만, 통계적으로 유의하지는 않는 것으로 나타났다.

대상가구원수 대비 응답가구원수를 살펴보기에 앞서, 조사 성공의 정

〈표 7-2〉 조사도구별 조사성공률 차이

		조사시작 가구	조사성공 가구	조사 성공률	S.D	diff[1)]	t-value	Pr>|t|
가구	CAPI	178	151	0.8483	0.3597	0.2281	5.72	<0.0001
	CAWI	495	307	0.6202	0.4858			
개인	CAPI	326	322	0.9877	0.1103	0.0900	5.17	<0.0001
	CAWI	567	509	0.8977	0.3033			

주 : 1) diff는 CAPI 조사성공률에서 CAWI 조사성공률을 뺀 차이를 의미함.

〈표 7-3〉 조사도구별 접촉가구 대비 조사성공률 차이(가구 설문)

	접촉가구	조사완료 가구	조사 성공률	S.D	diff[1)]	t-value	Pr>|t|
CAPI	280	151	0.5393	0.4993	-0.0335	-0.91	0.3609
CAWI	536	307	0.5728	0.4951			

주 : 1) diff는 CAPI 조사성공률에서 CAWI 조사성공률을 뺀 차이를 의미함.

의에 주의할 필요가 있다. 본 실험에서 조사 성공의 목표는 해당 가구의 15세 이상 가구원 전수가 설문조사에 답하는 것이다. 그러나 대상가구원 전수 조사를 고집할 경우, 가구 탈락률이 매우 높을 것은 자명한 일이다. 노동패널 본조사에서도 면접원에게 대상가구원 모두를 조사하도록 강제하고, 대상가구원 전수를 못하는 경우 따로 리포트를 하도록 하여 관리하고 있지만, 매년 약 90% 정도의 응답률을 보이고 있다. 따라서 본 실험에서도 노동패널 본조사의 지침과 동일하게 모든 가구원이 설문에 응해야 함을 강조하고, 실제로 자료 분석 시 성공가구의 의미는 가구설문과 가구에 속한 개인설문이 1개 이상 응답되었을 경우로 정의하였다. 따라서 대상가구원 대비 응답가구원의 비중은 웹조사 도입에 매우 중요한 요인이 될 것이라 생각된다.

대상가구원수 대비 응답가구원수를 살펴보면, CAPI의 경우 대상가구원 2명 가구에서 1명만 응답한 4가구를 제외하고, 목표 대상인 15세 이상 가구원 전수를 조사하였다. 그러나 CAWI의 경우 109(전체 307가구 대비 36%, 음영부분)가구가 대상가구원 전원조사에 실패하였다.

가구별로 대상가구원수 대비 응답가구원수 응답비율을 따로 계산해 보았다. 예를 들어, 15세 이상 대상가구원이 2명인 가구에서 개인설문을 1명만 응답한 경우에는 응답률이 50%이다. CAPI는 가구별로 평균 98.7%의 응답률을 보이고, CAWI는 평균 81.2%만 응답하는 것으로 조사되었

〈표 7-4〉 대상가구원수 대비 응답가구원수

(단위 : 가구)

15세 이상 가구원수	응답가구원수											
	CAPI						CAWI					
	1명	2명	3명	4명	5명	합계	1명	2명	3명	4명	5명	합계
1명	31	0	0	0	0	31	66	0	0	0	0	66
2명	4	79	0	0	0	83	69	100	0	0	0	169
3명	0	0	22	0	0	22	15	2	18	0	0	35
4명	0	0	0	12	0	12	9	5	5	12	0	31
5명	0	0	0	0	3	3	2	0	1	1	2	6
합계	35	79	22	12	3	151	161	107	24	13	2	307

〈표 7-5〉 가구별 대상가구원수 대비 응답가구원수 비율 차이

	사례수	평균	S.D	diff[1)]	t-value	Pr>|t|
CAPI	151	0.9868	0.0806	0.1752	7.97	<0.0001
CAWI	307	0.8115	0.2641			

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

〈표 7-6〉 조사도구별 대상가구원 전원 조사 기준 조사성공률 차이

	조사시작 가구	대상가구원 전원조사완료	조사 완료율	S.D	diff1)	t-value	Pr>|t|
CAPI	178	147	0.8258	0.3803	0.4258	10.50	<0.0001
CAWI	495	198	0.4000	0.4904			

주 : 1) diff는 CAPI 조사완료율에서 CAWI 조사완료율을 뺀 차이를 의미함.

다. CAWI의 평균 응답률이 CAPI에 비해 현저히 낮은 것은 통계적으로 유의하였다. 즉 CAWI조사 도입 시에는 현재 노동패널 개인 응답자 숫자를 유지하기 위해서는 표집 가구의 숫자를 CAPI보다 많이 산정해야 함을 의미한다.

대상가구원 전원 응답한 가구만이 조사에 성공했다고 가정하고, 조사성공률을 다시 구해 보았다. 15세 이상 대상가구원 전수를 조사 완료한 가구는 CAPI 82.6%, CAWI 40.0%로, CAPI가 통계적으로도 유의하게 높게 나타났다.

나. 응답시간

가구설문 조사도구별 응답일수를 살펴보면, CAPI는 조사당일 설문이 완료된 경우가 124건(82%)으로 압도적으로 많은 반면, CAWI는 당일, 1~2일, 3~4일, 5일 이상이 골고루 분포되어 있다. 평균 응답일수도 CAPI는 1.4일, CAWI는 4.4일로 유의하게 오래 걸리는 것으로 조사되었다.

가구설문 응답시간은 CAPI 27.37분, CAWI 27.40분으로 조사되었고, 통계적으로는 유의하지 않았다. 반면 가구원설문 응답시간은 가구설문과 다른 결과를 보인다. CAPI 29.5분, CAWI 39.5분으로 통계적으로도 유의

〈표 7-7〉 조사도구별 응답일수(가구 설문)

	사례수	당일	1~2일	3~4일	5일 이상
CAPI	151	124	10	6	11
CAWI	307	64	61	90	92

〈표 7-8〉 조사도구별 응답일수 차이(가구 설문)

	사례수	평균(일)	S.D	diff[1)]	t-value	Pr>|t|
CAPI	151	1.38	4.9244	-3.0623	-5.79	<0.0001
CAWI	307	4.44	5.5021			

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

〈표 7-9〉 조사도구별 응답시간 차이

		사례수	평균(분)	S.D	diff[1)]	t-value	Pr>|t|
가구	CAPI	151	27.37	11.2020	-0.0226	-0.01	0.9911
	CAWI	495	27.40	23.6527			
가구원	CAPI	255	29.45	18.0775	-10.0552	-3.35	0.0008
	CAWI	509	39.51	46.1628			

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

하게 CAWI가 오래 걸리는 것으로 나타났다. 그러나 CAWI의 경우, 응답시간이 2시간 이상으로 조사된 응답자가 20명(약 4%)이어서 평균값을 상승시킨 것으로 생각된다. 실재로 중위값은 28.2분으로 조사되었다. 이러한 문제가 발생한 이유는 CAWI는 웹에 접속된 상태로 설문조사에 응답하는 것이 아니라 다른 일을 하여도 통제할 수 없어 서버에 접속된 시간을 모두 응답시간으로 잡기 때문이다.

다. 조사 독려 방법 및 횟수

조사도구별로 조사 성공을 위한 조사 독려는 다른 방법으로 실시되었다. CAPI는 일차로 전화로 방문 약속을 잡고, 직접 방문하여 설문을 진행하였기 때문에 방문횟수만을 살펴보았다. 첫 번째 방문에서 조사 완료된

〈표 7-10〉 조사도구별 연락 횟수(가구 설문)

	CAPI(n=178)			CAWI(n=495)			
	방문			이메일		전화	
횟수	성공	비성공	일시	성공	비성공	성공	비성공
1회	121	27	10월 18일	30	0		
2회	22	0	10월 25일	237	45		
3회	8	0	11월 2일	36	114		
4회			11월 11일	3	23	15	133
5회			11월 16일	1	6	2	40
			11월 23일			0	15
전체	151	27	전체	307	188	17	188
평균(회)	1.25	1.00	평균(회)	2.05	2.95	1.12	1.37

가구는 121가구(80.1%)였다. CAWI는 일차로 10월 18일에 이메일로 설문을 발송하였고, 바로 조사 완료한 가구는 30가구(9.8%)로 CAPI에 비해 현저히 낮으며, 일주일 뒤에 발송된 두 번째 이메일 이후에 조사 완료한 가구는 237가구(77.2%, 누적 87.0%)에 달했다. CAWI의 전화 연락은 3차 이메일 독려 이후에 진행되었는데, 전화독려를 통해 조사에 성공한 가구는 총 17건(전체 성공가구 307건 중 5.5%)이다. 즉 이메일 독려만 진행했을 때, 성공하지 못했을 가구를 성공시킨 비율이 5.5%로, 마지막 성공률을 높이기 위해서 사용할 수 있는 방법이라 생각된다.

즉 대면조사인 CAPI 조사는 첫 번째 방문에서 80%의 가구가 조사를 완료하였고, 웹을 기반으로 한 CAWI 조사는 두 번째 안내 이메일에서 총 87%의 가구가 조사를 완료하는 것으로 나타났다.

라. 대리응답

웹조사의 주요한 단점은 본인이 직접 설문에 응답하였는지 확인할 수 없다는 점이다. 대리응답 여부는 CAPI 0.4%, CAWI 10.8%로 CAWI가 현저히 많았고, 통계적으로 유의하였다. 매년 노동패널조사에서 5% 미만의 대리응답이 존재하는 것을 감안할 때, CAWI의 대리응답 수치는

〈표 7-11〉 조사도구별 대리응답 분포 차이

	본인	대리	본인+대리	전체	Pr<=P
CAPI	248(97.3)	1(0.4)	6(2.4)	255	<0.0001
CAWI	422(82.9)	55(10.8)	32(6.3)	509	
전체	670(87.7)	56(7.3)	38(5.0)	764	

주 : Fisher's exact test 사용.

높은 것이라 생각된다.

대리응답 사유로 가장 많은 이유가 '컴퓨터를 다룰지 몰라서'가 11건으로 가장 많아, 가구원 전원을 웹조사로 진행하는 것은 컴퓨터를 사용할 줄 모르는 사람은 체계적으로 대리응답이나 설문에서 탈락될 확률이 생길 수 있음에 유념해야 한다. 또한 대리응답 사유로 주목할 만한 다른 사유는 '군복무로 할 수가 없어서', '멀리 떨어져 있어서', '같이 거주하지 않고 타 지방에 거주함' 등 분가의 의미나 비동거의 의미를 제대로 파악하지 못해서 발생한 4건이다. 즉 설문을 하지 않아도 되는 대상자까지 대상자에 포함시키고, 이를 대리 응답한 것이다. 건수는 많지 않지만 면접원이 없이 컴퓨터 화면에 보이는 설명만으로는 설문의 의미를 제대로 이해하지 못할 수도 있다는 점을 기억해야 할 것이다.

마. 조사도구별 개인 특성

조사도구별 응답자의 성별과 교육연수 차이를 살펴보면, 통계적으로 유의한 차이가 없었다.

〈표 7-12〉 조사도구별, 성별 교육연수 차이

		사례수	평균	S.D	diff[1)]	t-value	Pr>|t|
성별	CAPI	322	0.521739	0.500305	-0.03818	-1.0763	0.2821
	CAWI	509	0.559921	0.496885			
교육	CAPI	322	14.63665	2.471323	0.18478	1.0393	0.299
	CAWI	509	14.45187	2.512986			

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

〈표 7-13〉 조사도구별 연령 차이

	사례수	연령	S.D	diff[1)]	t-value	Pr>|t|
CAPI	322	42.63975	14.25777	1.826392	1.9171	0.0556
CAWI	509	40.81336	12.79358			

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

조사도구별로 연령 차이를 살펴보면, CAPI가 평균 42.6세, CAWI가 40.8세로 약간 낮았고, 10%의 유의수준에서 통계적으로 유의한 차이가 있는 것으로 조사되었다. 즉 컴퓨터를 이용하는 조사에서 더 연령이 낮은 것으로 나타나, CAWI 조사를 진행하는 경우, 고연령층의 대상자가 체계적으로 배제될 수 있음을 고려해야 할 것이다.

2. 항목응답률

가. 가구소득 변수

가구설문에서 주요한 항목 중의 하나인 가구소득에 대해 살펴보았다. 가구소득은 크게 근로소득, 금융소득, 부동산소득, 사회보험소득, 이전소득 및 기타소득으로 구성(총 6가지)되어 있다. 작년 한 해 동안의 가구소득과 지난 한 달 동안의 가구소득을 질문하며, 근로소득을 제외하고 나머지 소득 문항들은 각각 서브문항(subquestion)이 있어서 구체적으로 조사하고 있다. 본 보고서에서는 세부 항목을 합하여 여섯 가지 소득별로 항목응답률이 CAPI와 CAWI 간 차이가 있는지 분석하였다. 참고로 사회보험 소득액은 CAPI와 CAWI에서 모두 수혜받고 있지 않다고 응답하여 분석에서 제외하였다.

〈표 7-14〉 가구 근로소득액 응답률(CAPI vs CAWI)

	CAPI	CAWI
작년 한 해 근로소득	92.1	86.6
지난 한 달 근로소득	86.1	85.0

〈표 7-15〉 가구소득 응답률의 집단(CAPI vs CAWI) 간 차이

	CAPI		CAWI		diff[1]	t-value	Pr>|t|
	평균	S.D	평균	S.D			
소득총액[2]	0.9735	0.1611	0.9609	0.1941	0.0126	0.73	0.4635
근로소득액	0.9205	0.2714	0.8664	0.3407	0.0541	1.84	0.0669
금융소득액	0.3510	0.4789	0.4658	0.4996	-0.1148	-2.34	0.0195
부동산소득액	0.1722	0.3788	0.0912	0.2884	0.081	2.32	0.0213
이전소득액	0.4901	0.5016	0.2541	0.4360	0.2360	4.94	<.0001
기타소득액	0.2980	0.4589	0.3225	0.4682	-0.0245	-0.53	0.5970

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.
2) 소득총액은 근로소득액, 금융소득액, 부동산소득액, 이전소득액과 기타소득액을 모두 합한 값임.

<표 7-14>를 보면 작년 한 해 근로소득액의 응답률이 CAPI는 92.1%이고 CAWI는 86.6%로, 지난 한 달 근로소득액 역시 CAPI가 CAWI보다 1.1% 정도 높은 것으로 나타나 가구 근로소득액의 응답률은 모두 CAPI가 높은 편으로 나타났다.

가구소득 응답률이 조사도구 간 차이가 있는지를 분석한 결과 금융소득액과 기타소득액을 제외하고는 모두 CAPI에서 응답률이 높게 나타났다. 금융소득액과 부동산소득액 및 이전소득액은 통계적으로 유의한 차이가 있는 것으로 분석됐다. 또한, 근로소득액도 유의수준 10%에서는 통계적으로 유의한 차이가 있어서 기타소득액을 제외하고는 응답률이 조사도구 간 차이가 있다고 볼 수 있다. 근로소득액, 금융소득액, 부동산소득액, 이전소득액과 기타소득액을 모두 합한 값인 생성변수 소득총액은 CAPI의 응답률이 CAWI보다 1.26% 높은 97.35%로 거의 차이가 없으며, t-test 결과 역시 통계적으로 유의한 차이가 없는 것으로 나타났다.

나. 금융자산 변수

금융자산은 6개의 서브문항으로 구성되어 있으며 분석 시에는 하나의 변수로 통합하였다. 금융자산 역시 모든 항목에서 무응답이 관측되지 않았으며 응답률의 집단(CAPI vs CAWI) 간 차이를 분석한 결과, 응답률은

〈표 7-16〉 금융자산 응답률의 집단(CAPI vs CAWI) 간 차이

	CAPI		CAWI		diff[1)]	t-value	Pr>|t|
	평균	S.D	평균	S.D			
금융자산	0.9404	0.2375	0.8697	0.3372	0.0707	2.59	0.0099

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

집단 간 통계적으로 유의한 차이가 있는 것으로 나타났다. CAPI의 응답률이 CAWI보다 7.43% 정도 높다.

다. 지난 한 달 소득 변수

지난 한 달 소득은 6개의 서브문항을 가지고 있으며 분석 시 하나의 변수로 통합하였다. 모든 항목에서 무응답이 관측되지 않았다. 응답률 집단(CAPI vs CAWI) 간 차이를 분석한 결과, 응답률은 집단 간 통계적으로 유의한 차이가 있는 것으로 나타났으며 CAPI가 CAWI보다 응답률이 6.46% 정도 높다.

〈표 7-17〉 지난 한 달 소득 응답률의 집단(CAPI vs CAWI) 간 차이

	CAPI		CAWI		diff[1)]	t-value	Pr>|t|
	평균	S.D	평균	S.D			
지난 한 달 소득	0.9669	0.1795	0.9023	0.2974	0.0646	2.88	0.0041

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

라. 월평균 저축액

월평균 저축액은 저축액 전체를 묻는 항목과 8개의 서브문항을 묻는 항목으로 구성되어 있다. 저축액 전체는 8개의 서브문항을 합한 값과 동일하므로 분석 결과는 동일하다. 여기서는 저축액 전체를 묻는 항목으로 분석을 실시하였다. 모든 항목에서 무응답이 관측되지 않았으며 응답률의 집단(CAPI vs CAWI) 간 차이를 분석한 결과는 유의수준 10%에서는 통계적으로 유의한 차이가 있는 것으로 나타났다. 응답률은 CAPI가

〈표 7-18〉 월평균 저축액 응답률의 집단(CAPI vs CAWI) 간 차이

	CAPI		CAWI		diff[1)]	t-value	Pr>|t|
	평균	S.D	평균	S.D			
월평균 저축액	0.9007	0.3001	0.8404	0.3668	0.0603	1.87	0.0618

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

CAWI보다 6.03% 정도 높게 나타났다.

마. 생활비 및 주거 관련 변수

가구 생활비와 주거 관련 변수에 대해 살펴보았다. 가구 생활비는 작년 한 해 월평균 생활비를 전체 묻는 항목과 20개의 서브문항을 묻는 항목으로 구성되어 있다. 주거 관련 변수는 입주형태와 주택종류에 따라 대지면적과 평수를, 시가와 임대보증금, 월세금 등을 질문하고 있다. 여기서는 가구 생활비는 전체와 생활비 합계로 생성된 변수를 보았고 주거 관련 변수는 자가와 전월세로 구분하여 분석하였다.

가구 생활비를 살펴보면 전체와 생활비 합계 모두 응답률의 집단 간 차이는 통계적으로 유의한 차이가 없는 것으로 나타났다. 응답률은 CAPI는 100%로 전체 응답하였으나 CAWI는 헌금 및 각종 기부금에서 무응답이 있어 전체는 99.35%이고 생활비 합계에서는 99.67%였다. 무응답과 관련된 내용은 4항(항목무응답)에서 다른 문항들과 함께 분석하였다.

주거지 정보 응답률의 집단 간 차이는 모두 통계적으로 유의한 차이가 없는 것으로 나타났다. 응답률은 자가인 경우 CAPI가 높았고, 전월세의 경우에는 CAWI가 높게 나왔다.

〈표 7-19〉 가구 생활비 응답률의 집단(CAPI vs CAWI) 간 차이

	CAPI		CAWI		diff[1)]	t-value	Pr>|t|
	평균	S.D	평균	S.D			
전체	1	0	0.9935	0.0806	0.00651	1.42	0.1576
생활비 합계	1	0	0.9967	0.0571	0.00326	1.00	0.3181

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

〈표 7-20〉 주거지 정보 응답률의 집단(CAPI vs CAWI) 간 차이

	CAPI		CAWI		diff[1)]	t-value	Pr>|t|
	평균	S.D	평균	S.D			
자가 : 평수	0.5364	0.5003	0.5049	0.5008	0.0315	0.63	0.5265
자가 : 평당시가	0.5960	0.4923	0.5570	0.4976	0.0390	0.79	0.4289
전월세 : 평수	0.4040	0.4923	0.4430	0.4976	-0.0390	-0.79	0.4289
전월세 : 임대보증금	0.4040	0.4923	0.4430	0.4976	-0.0390	-0.79	0.4289
전월세 : 월세	0.4040	0.4923	0.4430	0.4976	-0.0390	-0.79	0.4289

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.
2) 자가에서 단독주택의 경우는 평수 대신 대지면적과 연건평을 질문하여 같이 분석할 수 없으므로 제외함.

바. 일자리 관련 변수

여기서는 일자리 관련 변수 중 평균소득 유무와 종업원수 응답률의 조사방법에 따른 차이를 분석한다. 분석에 앞서 조사방법에 따른 임금근로자 비중 차이를 살펴볼 필요가 있을 것이다. <표 7-21>에서 조사방법에 따른 임금근로자 비중 차이가 유의하지 않음을 알 수 있다.

<표 7-22>는 조사방법에 따른 종업원수 응답률 차이를 비교한 것이다. 조사방법에 따른 응답률 차이가 임금근로자와 비임금근로자 모두 유의

〈표 7-21〉 임금근로자 비중 조사방법에 따른 차이 비교

CAPI		CAWI		diff[1)]	t-value	Pr>|t|
임금근로자 비중	S.D.	임금근로자 비중	S.D.			
81.9%	0.3859	87.1%	0.3359	-5.2%p	1.52	0.1307

주 : 1) diff는 CAPI와 CAWI의 임금근로자 비중 차이를 의미.

〈표 7-22〉 종업원수 응답률의 조사방법에 따른 차이 비교

	CAPI		CAWI		diff[1)]	t-value	Pr>|t|
	응답률	S.D.	응답률	S.D.			
임금근로자	84.1%	0.3666	81.6%	0.3880	2.5%p	0.66	0.5108
비임금근로자	37.5%	0.4919	41.3%	0.4978	-3.8%p	-0.33	0.7396

주 : 1) diff는 CAPI와 CAWI의 응답률 차이를 의미.

하지 않음을 알 수 있다.

사. 흡연 및 음주

개인의 성향, 의견, 개인정보 등과 같은 응답자가 응답하기 꺼리는 민감한 질문에 대한 응답비중으로 집단 간 차이가 있는지 살펴보았다. 우선 <표 7-23>에서는 흡연 및 음주 항목에서 무응답이 관측되지 않았으나 응답률('있음'에 대한 응답비중)의 집단 간 차이를 분석한 결과, 음주 및 금주 변수의 응답률은 5% 수준에서 집단 간 통계적으로 유의한 차이가 있는 것으로 나타났다. 흡연 항목은 10% 수준에서 유의한 차이가 있는 것으로 보인다.

〈표 7-23〉 흡연 및 음주 응답비중의 집단 간 차이

	CAPI		CAWI		diff[1)]	t-value	Pr>|t|
	평균	S.D.	평균	S.D.			
흡연 여부	0.1294	0.3363	0.1807	0.3852	-0.0513	-1.89	0.0588
금연 시도	0.2121	0.4151	0.5109	0.5026	-0.2987	-3.06	0.0027
음주 여부	0.6588	0.4750	0.6267	0.4842	0.0321	0.87	0.3847
금주 시도	0.9804	0.1389	0.9528	0.2122	0.0275	2.15	0.0319

주 : 1) diff는 CAPI와 CAWI의 응답률 차이를 의미.

3. 분포의 차이

가. 직업훈련, 사회보험 및 혼인상태 변화

<표 7-24>는 개인적 민감 문항인 직업훈련 경험 여부, 자격증 보유 여부, 사회보험 수급 여부, 혼인상태 변화 여부 항목에 대한 응답률이다. 4개 항목에서 무응답은 관측되지 않았으며 응답률('있음'에 대한 응답비중)의 집단 간 차이를 분석한 결과, 직업훈련 경험과 사회보험 급여 수급 변수의 응답률은 집단 간 통계적으로 유의한 차이가 있는 것으로 보인다. 자격증 보유 항목은 통계적으로 유의하지 않은 것으로 나타났으며, 혼인

상태 변화 여부 항목의 집단 간 응답률을 비교하였지만 통계적으로 차이가 없는 것으로 나타났다.

CAPI와 CAWI의 차이가 응답 값의 분포에 따라 차이가 있는지를 보기 위하여 가구설문에서는 입주형태, 주택형태, 가구원수, 소득, 금융자산 및 저축액 문항을 통해 살펴보았다.

<표 7-25>는 입주형태 및 가구원수는 통계적으로 유의한 차이가 없음을 확인할 수 있었으나, 주택형태는 통계적으로 유의한 차이를 보였다.

〈표 7-24〉 민감 문항 응답률의 집단 간 차이

	CAPI		CAWI		diff[1)]	t-value	Pr>|t|
	평균	S.D.	평균	S.D.			
직업훈련 경험	0.1255	0.3319	0.2554	0.4365	-0.1299	-4.57	<.0001
자격증 보유	0.4902	0.5009	0.4401	0.4969	0.0501	1.31	0.1902
사회보험 수급	0.0980	0.2980	0.1807	0.3852	-0.0827	-3.27	0.0011
혼인상태 변화	0.0777	0.2683	0.1079	0.3107	-0.0302	-1.23	0.2191

주 : 1) diff는 CAPI와 CAWI의 응답률 차이를 의미.

〈표 7-25〉 입주 · 주거형태 및 가구원수 차이 검증

	CAPI : CAWI
입주형태	0.6850
주택형태	0.0164
	평균차 검정
가구원수	0.8972

주 : 1) 검정의 p값을 보고한 것임. 입주형태, 주택형태는 분할표(contingency table) 독립성 검증(χ^2) p값 결과임.
2) 입주형태는 자가/전월세/기타, 주택형태는 단독/아파트/연립 · 다세대/상가/기타임.

나. 가구 소득 및 금융자산

소득 변수 차이를 검증한 결과, 근로소득액을 제외한 모든 항목(금융소득액, 부동산소득액, 이전소득액, 기타소득액)의 평균은 CAPI와 CAWI 간

통계적으로 유의한 차이가 없는 것으로 나타났다. 근로소득액의 평균은 CAPI는 5,672만 원 정도이고 CAWI는 4,903만 원 정도로 CAPI가 769.5만 원 정도 더 높게 나타났으며 t-test 결과는 통계적으로 유의한 차이가 있는 것으로 나타났다. 전반적으로 CAPI에서 소득 변수들의 값이 더 넓게 퍼져 있어서 표준편차의 값도 크다는 것을 알 수 있다. 비모수 검정에서는 유의수준 5%에서 유의한 차이가 없다는 것으로 나타났다. 참고로, 이상치로 판단되는 케이스를 제외하고 분석해 보았는데 동일한 결과를 보였다. 그러나 이상치로 판단할 수 있는 근거가 뚜렷하지 않으므로 값이 있는 경우는 모두 포함하여 분석하였다. 다음으로 분산에는 차이가 있는지를 확인하기 위해 분포형태와 관계없이 검정할 수 있는 Levene 검정도 실시하였다. 소득 변수들 모두 CAPI와 CAWI 간의 분산은 통계적으로 유의한 차이가 없음을 확인하였다.

추가적으로 금융자산, 지난 한 달 소득과 월평균 저축액의 분포에 대해서도 살펴보았다. 우선 금융자산의 평균은 집단 간 통계적으로 차이가 없는 것으로 나타났다. 그러나 비모수 검정에서는 유의수준 10%에서 통계적으로 유의한 차이가 있다. 지난 한 달 소득의 평균 및 월평균 저축액 모두 집단 간 통계적으로 차이가 없는 것으로 나타났다. 그러나 두 변수 모두 비모수 검정에서는 유의수준 5%에서 통계적으로 유의한 차이가 있는 것으로 나타났다. 금융자산, 지난 한 달 소득 및 월평균 저축액 항목 모두

〈표 7-26〉 소득 변수 차이 검증

	CAPI		CAWI		diff[1]	t-value	Pr>|t|	Pr>|Z|[2]	Levene 검정
	평균 (만 원)	S.D	평균 (만 원)	S.D					
근로소득액	5,672.3	3,731.5	4,902.8	2,831.2	769.5	2.13	0.0341	0.0526	0.1054
금융소득액	308.7	547.9	1,156.4	6,195.8	-847.6	-1.62	0.1076	0.9627	0.5248
부동산소득액	1,614.2	1,574.1	4,966.4	16,373.7	-3,352.2	-1.08	0.2904	0.9310	0.3166
이전소득액	1,683.9	4,889.3	5,118.4	33,987.0	-3,434.5	-0.88	0.3799	0.5197	0.3308
기타소득액	1,235.2	3,916.8	4,764.0	36,692.2	-3528.8	-0.95	0.3468	0.7026	0.4998

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.
2) Pr>|Z|는 비모수 검정(nonparametric test)을 통해 계산된 값임.

〈표 7-27〉 금융자산 및 기타변수 차이 검증

	CAPI		CAWI		diff[1]	t-value	Pr>|t|	Pr>|Z|[2]	Levene 검정
	평균 (만 원)	S.D	평균 (만 원)	S.D					
금융자산	7,406.5	16,901.7	6,415.9	11,865.2	990.6	0.62	0.5348	0.0861	0.3357
지난 한 달 소득	561.2	928.6	461.0	674.9	100.3	1.15	0.2497	0.0433	0.4651
월평균 저축액	111.8	97.0	102.1	164.7	9.7	0.73	0.4629	0.0184	0.2594

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.
2) Pr>|Z|는 비모수 검정(nonparametric test)을 통해 계산된 값임.

CAPI와 CAWI 간의 분산은 통계적으로 유의한 차이가 없는 것으로 나타났다.

다. 만족도 및 건강상태

개인의 주관적 만족도 항목인 조직몰입도, 직무만족도, 요인별 직무만족도, 생활만족도, 건강상태 관련 변수의 집단 간 응답 차이를 분석해 보았다. 만족도 문항은 리커트 5점 척도에 의해 조사되며, 분석을 위해 두 집단 평균차 검정을 수행하였다.

<표 7-28>에서 보듯이 직장만족도 전체 21개 항목 모두 CAPI와 CAWI 두 집단 간 평균점수 차이가 5% 수준에서 통계적으로 유의하게 나타났다. 조직몰입도 5개 변수와 직무만족도 5개 변수의 CAWI 집단의 평균점수는 CAPI 집단의 평균점수보다 통계적으로 유의하게 낮은 것으로 분석되었다. 주관적 만족도 변수에서 CAPI와 CAWI 조사방식의 변화로 인한 모드효과[45]가 존재한다고 볼 수 있을 것 같다.

<표 7-29>에서는 생활만족도 차이를 검증해 보았다. 생활만족도 7개 항목에서 CAWI 집단의 평균점수는 CAPI 집단의 평균점수보다 통계적으로 유의하게 높은 것으로 분석되었다.

45) 모드효과란 조사방식(면접조사, 전화조사, 우편조사, 인터넷조사 등)의 변화로 인해서 발생하는 표본이탈, 무응답오차, 측정오차 등을 종합적으로 일컫는 용어(최효미 외(2012), 『패널자료의 품질 개선을 위한 조사기법 연구』, p.27 참조).

〈표 7-28〉 직장만족도 응답 차이 검증

		CAPI : CAWI		
		CAPI	CAWI	p값
조직 몰입도	지금 근무하고 있는 직장은 다닐 만한 좋은 직장이다.	3.6690	3.4903	0.0397
	나는 이 직장에 들어온 것을 기쁘게 생각한다.	3.6621	3.4355	0.0093
	직장을 찾고 있는 친구가 있으면 나는 이 직장을 추천하고 싶다.	3.4414	3.1548	0.0043
	나는 내가 다니고 있는 직장을 다른 사람들에게 자랑할 수 있다.	3.5517	3.2581	0.0021
	별다른 일이 없는 한 이 직장을 계속 다니고 싶다.	3.7793	3.5581	0.0158
직무 만족도	나는 현재 하고 있는 일에 만족하고 있다.	3.7062	3.3127	<.0001
	나는 현재 하고 있는 일을 열정적으로 하고 있다.	3.6441	3.3634	0.0002
	나는 현재 하고 있는 일을 즐겁게 하고 있다.	3.6836	3.2789	<.0001
	나는 현재 하고 있는 일을 보람을 느끼면서 하고 있다.	3.6836	3.3155	<.0001
	별다른 일이 없는 한 현재 하고 있는 일을 계속 하고 싶다.	3.8531	3.4085	<.0001
요인별 직무 만족도	임금/보수 만족도	2.8588	3.0394	0.0226
	취업의 안정성 만족도	2.4124	2.7549	<.0001
	하고 있는 일의 내용 만족도	2.2994	2.6732	<.0001
	근무환경 만족도	2.3842	2.7324	<.0001
	근로시간 만족도	2.3898	2.7380	<.0001
	개인의 발전가능성 만족도	2.5141	2.9070	<.0001
	의사소통 및 인간관계 만족도	2.3277	2.7465	<.0001
	인사고과의 공정성 만족도	2.6483	2.9774	0.0002
	복지 후생제도 만족도	2.7931	3.1677	0.0002
전반적 일자리(직장) 만족도		2.4068	2.7775	<.0001
전반적 일(직무) 만족도		2.3390	2.7268	<.0001

주 : 조직몰입도, 직무만족도는 (1)매우 불만족~(5)매우 만족, 요인별 직무만족도와 전반적 만족도는 (1)매우 만족~(5)매우 불만족이므로 주의.

<표 7-30>은 건강상태 변수의 차이를 검증하고 있다. 현재 건강상태, 다른 사람 대비 건강상태 항목은 CAPI 집단보다 CAWI 집단의 평균점수가 높은 것으로 나타났으며, 통계적으로 유의한 결과를 보였다. 또한 감

〈표 7-29〉 생활만족도 차이 검증

	CAPI : CAWI		
	CAPI	CAWI	p값
가족의 수입	2.9098	3.1650	<.0001
여가활동	2.8078	3.1022	<.0001
주거환경	2.5804	2.9627	<.0001
가족관계	2.2118	2.4951	<.0001
친인척관계	2.5176	2.8762	<.0001
사회적 친분관계	2.3686	2.7348	<.0001
전반적 생활만족도	2.5294	2.8251	<.0001

주 : 생활만족도는 (1)매우 만족~(5)매우 불만족인 5점 척도.

〈표 7-30〉 건강상태 및 장애 응답 차이 검증

	CAPI : CAWI		
	CAPI	CAWI	p값
현재 건강상태	2.3608	2.6346	<.0001
1년 전 대비 건강상태	2.9490	3.0275	0.1249
다른 사람 대비 건강상태	2.5529	2.7937	<.0001
감각기관 장애	1.9882	1.9587	0.0082
육체적 제약	1.9843	1.8998	<.0001
활동제약 1 (배우기, 기억하기, 집중하기)	1.9686	1.9214	0.0037
활동제약 2 (옷입기, 목욕하기 등)	1.9882	1.9764	0.2163
활동제약 3 (쇼핑하기, 병원가기 등)	1.9882	1.9725	0.1131
활동제약 4 (직업활동 어려움)	1.9608	1.9077	0.0028

주 : 건강상태는 (1)매우 건강~(5)매우 불건강인 5점 척도. 장애 및 활동제약 항목은 (1)예 (2)아니오임.

〈표 7-31〉 흡연 및 음주 변수의 차이 검증

	CAPI		CAWI		diff[1]	t-value	Pr>|t|
	평균	S.D.	평균	S.D.			
흡연 여부[2]	2.5843	0.7097	2.4067	0.7772	0.1776	3.07	0.0023
하루 평균 흡연량	3.3030	0.5855	3.2065	0.6383	0.0965	0.76	0.4481
금연 시도	1.7879	0.4151	1.4891	0.5026	0.2987	3.06	0.0027
음주 여부	1.5725	0.8428	1.5894	0.8219	-0.0168	-0.26	0.7912
음주 빈도	2.4405	1.0012	2.4639	0.9926	-0.0235	-0.25	0.8048
금주 시도	1.9702	0.1704	1.9248	0.2642	0.0455	2.30	0.0220

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.
2) 흡연 여부는 (1)피운다 (2)과거에는 피웠으나 현재에는 피우지 않는다 (3)피워본 적이 없다; 음주 여부는 (1)마신다 (2)과거에는 마셨으나 현재는 마시지 않는다 (3)마셔본 적이 없다.

각기관 장애, 육체적 제약, 활동제약1(배우기, 기억하기, 집중하기), 활동제약4(작업활동 어려움)는 CAWI 집단보다 CAPI 집단의 평균점수가 높게 나타나 통계적으로 유의한 것으로 분석되었다.

<표 7-31>에서는 건강 관련 변수로 흡연과 음주 변수들을 비교해 보았다. 흡연 여부는 '흡연/비흡연/과거에 흡연한 적 있다'로 응답하게 되어 있는 변수이며, 하루 평균 흡연량은 개수를 연속형으로 질문한 것이 아니라 구간을 나누어 질문하였다. 금연 시도는 '시도한 적 있다/없다'로 구분해 답하도록 되어 있다. 음주 관련 질문들도 흡연과 유사하게 설계되어 있다. '흡연 여부', '금연 시도', '금주 시도' 등 3개 변수는 5% 수준에서 통계적으로 유의한 차이가 발견되었다.

라. 일자리 변수

여기서는 종사상 지위, 시간제 및 자기선언적 비정규직 근로자 여부, 종업원수, 평균근로시간 등 일자리 관련 변수의 조사방법에 따른 평균 및 분포 차이를 비교한다. <표 7-32>는 조사방법에 따른 종사상 지위 분포 차이를 비교한 것이다. 일부 종사상 지위의 표본수가 충분치 않아 분포 차이의 검정은 Fisher의 정확한 검정(Fisher's exact test)으로 이루어졌다. 검정 결과, 임금근로자와 비임금근로자 모두 종사상 지위 분포의 조사방법에 따른 차이가 유의하지 않은 것으로 나타났다.

〈표 7-32〉 종사상 지위 분포의 조사방법에 따른 차이 비교

	종사상 지위	CAPI		CAWI		Pr<=P
		표본수	백분율	표본수	백분율	
임금 근로자	상용직	126	86.9	287	92.6	0.1428
	임시직	15	10.3	18	5.8	
	일용직	4	2.8	5	1.6	
	전체	145	100.0	310	100.0	
비임금 근로자	고용주	11	34.4	19	41.3	0.8403
	자영업자	20	62.5	25	55.6	
	무급가족 종사자	1	3.1	2	4.4	
	전체	32	100.0	46	100.0	

<표 7-33>은 임금근로자 중 비정규직 근로자 비중의 조사방법에 따른 차이를 비교한 것이다. 비정규직 비중은 통계 기준과 자기선언 기준에 따라 각각 계산되었다. 여기서 통계 기준은 2002년 노사정위원회에서 합의한 바에 따른 것이며, 자기선언 기준은 응답자의 주관적 판단에 따른 것이다. 자기선언적 비정규직 근로자 비중은 응답을 거부한 표본을 제외하고 계산하였다.

CAWI의 통계 기준 비정규직 비율은 64.5%로 CAPI의 2배에 달하는 것으로 나타났다. 이러한 차이는 한시적근로자와 특수형태근로종사자의 비중 차이에 기인한다. 한시적근로자 비중은 CAPI에서 28.3%, CAWI에서 55.8%인 것으로 나타났고, 특수형태근로종사자는 각각 9.0%와 16.8%로 나타났다. 자기선언 기준 비정규직 비율은 반대로 CAPI(21.5%)가 CAWI(14.3%)보다 높았으나 그 유의성은 분명하지 않았다.

한시적근로자 비중의 조사방법에 따른 차이는 근로기간의 정함 여부에 대한 응답자들의 오해에 의한 것으로 생각된다. 근로기간의 정함 여부를 묻는 문항에 대해 CAPI에서는 11.3%가 정해져 있다고 응답한 반면 CAWI에서는 37.4%가 그렇다고 응답했다. 또한 조사된 근로기간 역시 두 조사방법 간 차이가 큰 것으로 나타났다. 근로기간이 정해져 있는 경우 그 근로기간은 CAPI에서는 0~2년이 대부분이었던 반면에 CAWI에서는 10년 이상의 장기인 경우가 비교적 많았고, 특히 20년으로 응답한 비율이

21.1%에 달했다. 이는 응답자들이 정년과 근로기간의 정함 여부의 차이에 대해 오해하고 있을 가능성을 시사한다. 특수형태근로종사자 비중의 차이는 독립도급업자 비중 차이에서 기인한다. 독립도급업자 비중은 CAPI에서는 4.0%였으나, CAWI에서는 13.0%에 달했다.

종사상 지위, 시간제 및 자기선언적 비정규직 근로자 여부 등은 당초 CAPI에서는 응답자가 관련 용어를 이해하지 못할 경우 조사자가 이를 설명할 수 있는 반면 CAWI에서는 그렇지 않다는 점이 우려되었으나, 조사방법에 따른 결과 차이는 임금근로자와 비임금근로자 모두 유의하지 않은 것으로 나타났다.

<표 7-34>는 주당 평균근로시간을 조사방법에 따라 비교한 것이다. 주당 평균노동시간에 대한 무응답은 두 조사방법 모두에서 발견되지 않았다. 임금근로자의 주당 평균근로시간은 정규근로시간이 정해져 있는 경우 정규근로시간과 초과근로시간의 합으로, 정해져 있지 않은 경우 주당 평균근로시간으로 측정하였다. 조사방법에 따른 주당 평균근로시간 차이는 유의하지 않은 것으로 나타났으며, 분산 차이 역시 유의하지 않은 것으로 나타났다.

〈표 7-33〉 임금근로자 중 비정규직 근로자 비중의 조사방법에 따른 차이 비교

	CAPI		CAWI		diff[1)]	t-value	Pr>|t|
	비중	S.D.	비중	S.D.			
통계 기준	0.3241	0.4697	0.6452	0.4792	-0.3210	-6.70	<.0001
자기선언기준	0.2153	0.4124	0.1433	0.3510	0.0720	1.81	0.0717
불일치여부[2)]	0.2069	0.4065	0.5194	0.5004	0.3125	-7.08	<.0001

주 : 1) diff는 CAPI와 CAWI의 시간제 및 비정규직 근로자 비중 차이를 의미.
2) 통계 기준과 자기선언 기준에 따른 비정규직 여부가 일치하지 않는 경우.

〈표 7-34〉 평균근로시간의 조사방법에 따른 차이 비교

	CAPI		CAWI		diff[1)]	t-value	Pr>|t|
	평균	S.D.	평균	S.D.			
임금근로자	43.9시간	12.47	45.3시간	12.65	-1.3	-1.04	0.2969
비임금근로자	44.0시간	17.69	43.5시간	15.84	0.5	0.14	0.8919

주 : 1) diff는 CAPI와 CAWI의 평균근로시간 차이를 의미.

마. 개인 평균소득

<표 7-35>는 평균소득 유무의 조사방법에 따른 차이를 보여준다. 비임금근로자 중 적자를 보고 있는 경우와 무급가족종사자는 평균소득이 없는 것으로 간주하였다. CAPI에서는 임금근로자 중 평균소득이 없다고 응답한 경우가 없었음에도 불구하고, 임금근로자와 비임금근로자 모두 평균소득 없음의 비중 차이가 유의하지 않은 것으로 나타났다.

<표 7-36>은 평균소득의 조사방법에 따른 차이를 비교한 것이다. 임금근로자의 평균소득은 월평균임금으로, 비임금근로자는 월평균소득으로 계산되었으며, 비임금근로자 중 적자를 보고 있는 경우와 무급가족종사자의 평균소득은 0으로 간주하였다. 모수 검정에서는 평균소득의 조사방법에 따른 차이는 유의하지 않은 것으로 나타났으나, 비모수 검정에서는 임금근로자의 평균소득 차이가 유의한 것으로 나타났다.

〈표 7-35〉 평균소득 유무의 조사방법에 따른 차이 비교

	CAPI		CAWI		diff[1)]	t-value	Pr>|t|
	평균소득 없음	S.D.	평균소득 없음	S.D.			
임금근로자	-	-	0.7%	0.0802	-0.7%p	-1.42	0.1576
비임금근로자	9.7%	0.3005	18.2%	0.3902	-8.5%p	-1.02	0.3118

주 : 1) diff는 CAPI와 CAWI의 평균소득 없음의 비중 차이를 의미.

〈표 7-36〉 평균소득의 조사방법에 따른 차이 비교

	CAPI		CAWI		diff[1)]	t-value	Pr>|t|
	평균소득	S.D.	평균소득	S.D.			
임금근로자	304.5만 원	194.6	309.7만 원	361.0	-5.2만 원	-0.20	0.8447
비임금근로자	254.5만 원	222.7	306.9만 원	382.5	-52.4만 원	-0.73	0.4650

주 : 1) diff는 CAPI와 CAWI의 평균소득 차이를 의미.

4. 항목무응답률

설문조사에 참여하였더라도 설문 항목에 응답을 하지 않아서 항목무응

답률이 높다면 신뢰성 있는 조사라고 생각하기 어려울 수 있다. CAPI는 응답자가 무응답을 하더라도 면접원이 응답을 하도록 유도하여 응답을 받아 낼 수 있는 데 반해, CAWI는 응답자가 혼자서 진행하기 때문에 응답하기 어렵거나 밝히고 싶지 않을 때 무응답으로 쉽게 넘어갈 수도 있다. 그래서 설문 화면에 모름/무응답 버튼을 넣어서 보여주었을 때 응답자가 모름/무응답 버튼을 사용할 가능성이 높은지에 대해 실험해 보았다.

본 실험에서는 모든 항목에 모름/무응답 버튼을 생성하기에는 예산상 시간상 한계가 있어 일부 주요항목들, 특히 숫자가 기록되는 변수로 무응답률이 높을 것이라 예상되는 변수에 버튼을 생성하였다. 변수 목록은 아래와 같다.

[가구]
h_2301 (작년한해) 월평균 생활비(만)
h_2311 (작년한해) 액수(만) - 식비
h_2312 (작년한해) 액수(만) - 외식비
h_2313 (작년한해) 액수(만) - 공교육비
h_2314 (작년한해) 액수(만) - 사교육비
h_2315 (작년한해) 액수(만) - 차량유지비
h_2316 (작년한해) 액수(만) - 주거비
h_2317 (작년한해) 액수(만) - 경조사비
h_2318 (작년한해) 액수(만) - 보건의료비
h_2319 (작년한해) 액수(만) - 문화비(→교양오락비)
h_2320 (작년한해) 액수(만) - 내구재
h_2321 (작년한해) 액수(만) - 통신비
h_2323 (작년한해) 액수(만) - 같이 사는 부모님 용돈
h_2324 (작년한해) 액수(만) - 자녀 용돈
h_2325 (작년한해) 액수(만) - 그 외 가구원의 용돈
h_2326 (작년한해) 액수(만) - 피복비
h_2327 (작년한해) 액수(만) - 헌금 및 각종 기부금
h_2328 (작년한해) 액수(만) - 국민연금
h_2332 (작년한해) 액수(만) - 건강보험료(의료보험 및 노인장기요양보험)
h_2329 (작년한해) 액수(만) - 대중교통비
h_2330 (작년한해) 액수(만) - 생필품 구입비
h_2331 (작년한해) 액수(만) - 기타
h_2602 (조사당시) 부채잔액(만) - 1.금융기관 부채
h_2603 (조사당시) 부채 상환금(월/만) - 1.금융기관 부채
h_2605 (조사당시) 부채잔액(만) - 2.비금융기관 부채
h_2606 (조사당시) 부채 상환금(월/만) - 2.비금융기관 부채

h__2608 (조사당시) 부채잔액(만) - 3.개인적으로 빌린 돈
h__2609 (조사당시) 부채 상환금(월/만) - 3.개인적으로 빌린 돈
h__2611 (조사당시) 부채잔액(만) - 4.전세금, 임대보증금 받은 것
h__2612 (조사당시) 부채 상환금(월/만) - 4.전세금, 임대보증금 받은 것
h__2614 (조사당시) 부채잔액(만) - 5.미리 타고 앞으로 부어야 할 계
h__2615 (조사당시) 부채 상환금(월/만) - 5.미리 타고 앞으로 부어야 할 계
h__2617 (조사당시) 부채잔액(만) - 6.기타
h__2618 (조사당시) 부채 상환금(월/만) - 6.기타

[개인]
p__1642 (주된일자리)임금근로자-월평균임금액수(만)
p__1672 (주된일자리)비임금근로자-월평균 소득(만)
p__1703 작년세후총연간근로소득(만)
p__1702 작년세후총연간근로소득(만)
p__0402 (주된일자리)전체종업원수(명)

[직업력]
j145 정규/비정규여부-현재(또는 최종)
j001 취업시기(년)
j002 취업시기(월)
j004 퇴직시기(년)
j005 퇴직시기(월)

가. 가구변수

가구설문에서는 월평균 가구 생활비와 부채 항목에서만 모름/무응답 버튼을 생성하였다. 모름/무응답 버튼을 보여주지 않는 가구설문의 다른 문항들에서는 전반적으로 무응답을 한 경우가 거의 없었다. 모름/무응답 버튼이 있는 가구 생활비와 부채는 무응답을 응답한 케이스가 발견되었으나 케이스 수는 많지 않았다. <표 7-37>을 보면 금융기관 부채의 잔액과 상환금은 CAPI의 무응답률이 더 높게 나왔으나, 나머지 항목들은 CAWI에서만 모름/무응답 버튼을 사용한 것으로 나타났다.

무응답률이 CAPI와 CAWI 간 차이가 있는지를 살펴보았다. 가구생활비는 통계적으로 유의한 차이가 있는 것으로, 부채 잔액 및 상환금은 통계적으로 유의한 차이가 없는 것으로 나타났다. 참고로, 부채 잔액 및 상

〈표 7-37〉 무응답이 있는 문항의 무응답 비율

항목	세부 항목		CAPI	CAWI
가구 생활비	헌금 및 각종 기부금	무응답	0 (0.00)	17 (12.88)
		전체	86	132
부채	금융기관(잔액)	무응답	2 (3.03)	3 (1.88)
		전체	66	160
	금융기관(상환금)	무응답	3 (4.55)	6 (3.75)
		전체	66	160
	비금융기관(상환금)	무응답	0 (0.00)	1 (14.29)
		전체	2	7
	개인적으로 빌린 돈(잔액)	무응답	0 (0.00)	1 (4.35)
		전체	4	23
	개인적으로 빌린 돈(상환금)	무응답	0 (0.00)	2 (8.70)
		전체	4	23
	전세금, 임대보증금 받은 것(잔액)	무응답	0 (0.00)	2 (5.88)
		전체	22	34

주 : 1) 괄호 안의 수치는 무응답률을 의미함.
2) 부채의 세부항목 기타는 무응답률이 0%여서 제시하지 않음.

〈표 7-38〉 가구 생활비 무응답률의 집단(CAPI vs CAWI) 간 차이

	CAPI		CAWI		diff[1)]	t-value	Pr>|t|
	평균	S.D	평균	S.D			
헌금 및 각종 기부금	0.0000	0.0000	0.0554	0.2291	-0.0554	-4.24	<.0001

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

〈표 7-39〉 부채(잔액과 상환금) 무응답률의 집단(CAPI vs CAWI) 간 차이

	CAPI		CAWI		diff[1)]	t-value	Pr>|t|
	평균	S.D	평균	S.D			
잔액	0.0247	0.1561	0.0267	0.1617	-0.00205	-0.1	0.9235
상환금	0.0429	0.204	0.0349	0.184	0.00797	0.3	0.7675

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

환금은 서브항목별 무응답 케이스가 작아서 하나의 변수로 통합하여 전체로 보았다. 조사도구와 상관없이 무응답률이 매우 낮은 것을 알 수 있

다. 추가로 독립성 검정을 실시해 본 결과 조사도구와 무응답 여부는 어떠한 연관성도 가지지 않는다는 것을 확인하였다.

나. 교육, 혼인상태 변수

개인설문에서 응답자가 응답하기 꺼리는 민감한 질문에 대한 무응답률을 살펴보고자 한다. 개인설문에서 모름/무응답 버튼을 허용한 작년 한 해 근로소득 변수를 제외하고 거의 대부분 변수는 무응답이 없는 것으로 확인되었다. 다만 혼인상태 문항은 로직 에러가 있는 문항으로 확인되었다.

<표 7-40>은 개인용 자료의 주요 변수 중 재학 중인 개인의 교육과 혼인상태에 관한 무응답률이 제시되어 있다. 신규 개인용 설문에서 다니셨거나 다니고 계신 고등학교/대학교/대학원에 대한 교육력 설문이 진행된다. 현재 재학 중인 학교의 유형에 대한 무응답은 나타나지 않았다.

한편 혼인력과 관련된 변수의 무응답률은 로직 에러로 인하여 무응답률을 확인할 수 있었다. 혼인상태가 미혼(한 번도 결혼한 적 없다)의 경우를 제외하고는 혼인상태 변화가 있었거나 없었던 개인에 대해 변화 내용

〈표 7-40〉 재학 여부, 혼인상태 무응답률(CAPI vs CAWI)

	CAPI		CAWI	
	무응답률		무응답률	
	리뷰 전	리뷰 후	리뷰 전	리뷰 후
재학 중인 학교의 유형	0.0	0.0	0.0	0.0
혼인상태	0.0	0.0	0.0	0.0
혼인상태 변화 여부	0.0	0.0	0.0	0.0
혼인상태 변화 횟수	0.0	0.0	0.0	0.0
혼인상태 변화 내용1	92.2	0.0	89.2	0.0
혼인상태 변화 내용2	0.0	0.0	0.0	0.0
혼인상태 변화 내용3	0.0	0.0	0.0	0.0
혼인상태 변화 내용4	0.0	0.0	0.0	0.0

주: 리뷰 후 자료는 기존 CAPI 버전의 프로그램을 CAWI(웹) 버전으로 변경하는 과정 중에 발생한 로직 에러로 인한 오류로 판단되어 자료 수정된 수치임.

-1에서 처음 결혼함으로 응답되어야 하나, 로직 에러로 인하여 혼인상태가 변화가 없는 개인에게 응답받지 못한 케이스를 발견하였다. 리뷰 후에는 해당 문항에 대한 응답 값이 전부 채워진 결과, 혼인상태 관련 변수의 무응답률은 발견되지 않았다.

다. 개인 근로소득 변수

<표 7-41>은 개인의 작년 한 해 근로소득 관련 문항의 t-test 결과이다. 세전/세후 근로소득 문항 t-test에 사용된 자료는 소득액이 있다는 응답자의 소득액 문항의 무응답 비중이다. 즉 무응답률의 결과는 문항에 대한 응답 빈도가 차이가 있는가를 검증한 것이다. t-test 결과, 세전 근로소득액과 세후 근로소득액은 조사도구에 따라 유의한 차이를 갖는 것으로 나타났다.

〈표 7-41〉 작년 한 해 근로소득 무응답률의 집단 간 차이

	CAPI		CAWI		diff[1)]	t-value	Pr > |t|
	평균	S.D	평균	S.D			
근로소득 유무	0.0000	0.0000	0.0000	0.0000	-	-	-
세전 근로소득	0.0118	0.1080	0.0943	0.2925	-0.0825	-5.64	<.0001
세후 근로소득	0.0235	0.1519	0.1238	0.3296	-0.1002	-5.75	<.0001

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미.

라. 일자리 관련 변수

여기서는 직업력 자료 중 응답자가 모름/무응답을 선택할 수 있는 변수들의 무응답률을 조사방법에 따라 비교한다. 직업력 자료 중 일자리 시작시기, 자기선언적 비정규직 근로자 여부, 종업원수, 평균소득은 응답자가 모름/무응답을 선택할 수 있다.

<표 7-42>는 일자리 시작시기의 무응답률을 조사방법에 따라 비교한 것이다. 검정 결과, 조사방법에 따른 일자리 시작시기 무응답률 차이는 유의하지 않았다. 무응답이 모두 비임금근로자에서만 나타났는데, 이는

실제 조사와 유사한 결과이다. 1~18차 한국노동패널조사에서 비임금근로자의 일자리 시작시기 무응답률(1.0%)은 임금근로자(0.3%)보다 높았다.

<표 7-43>은 임금근로자 중 자기선언적 비정규직 근로자 여부와 종업원수 무응답률을 조사방법에 따라 비교한 것이다. 비임금근로자의 종업원수 무응답률은 표본수가 적어 분석에서 제외하였다. 검정 결과, 조사방법에 따른 자기선언적 비정규직 근로자 무응답률 차이는 유의하지 않았다. 그러나 종업원수 무응답률은 유의한 차이를 보였다.

<표 7-44>는 평균소득 무응답률의 조사방법에 따른 차이를 비교한 것이다. 임금근로자의 경우 평균소득 무응답률의 유의한 차이가 발견되었다. 앞서 평균소득의 조사방법에 따른 차이가 유의하지 않았음을 볼 때, 이는 CAPI에서는 조사자가 대략적인 평균소득을 조사해 오는 반면에 CAWI에서는 응답자가 부정확한 답변을 피하려는 성향을 보이기 때문으로 생각된다. 보다 정확한 분석을 위해서는 자료의 히핑(heaping) 정도를

〈표 7-42〉 일자리 시작시기(연/월/일) 무응답률의 집단 간 비교

	일자리 시작시기	CAPI		CAWI		diff[1)]	t-value	Pr>|t|
		무응답	S.D.	무응답	S.D.			
임금 근로자	연	-	-	-	-	-	-	-
	월	-	-	-	-	-	-	-
	일	-	-	-	-	-	-	-
비임금 근로자	연	0.0313	0.1768	0.0217	0.1474	0.010	0.26	0.8036
	월	0.0313	0.1768	0.0217	0.1474	0.010	0.26	0.8036
	일	0.0313	0.1768	0.0217	0.1474	0.010	0.26	0.8036

주 : 1) diff는 CAPI와 CAWI의 무응답률 차이를 의미.

〈표 7-43〉 임금근로자 중 비정규직 여부와 종업원수 무응답률의 조사방법에 따른 차이 비교

	CAPI		CAWI		diff[1)]	t-value	Pr>|t|
	무응답	S.D.	무응답	S.D.			
비정규직 여부	0.7%	0.0830	1.0%	0.0981	0.3%p	-0.31	0.7539
종업원수	4.0%	0.1952	10.9%	0.3124	-7.0%p	-2.75	0.0062

주 : 1) diff는 CAPI와 CAWI의 무응답률 차이를 의미.

〈표 7-44〉 평균소득 무응답률의 조사방법에 따른 차이 비교

	CAPI		CAWI		diff[1]	t-value	Pr>|t|
	무응답	S.D.	무응답	S.D.			
임금근로자	0.7%	0.0830	5.2%	0.2216	4.5%p	-3.12	0.0020
비임금근로자	-	-	4.6%	0.2107	-4.6%p	-1.43	0.1597

주 : 1) diff는 CAPI와 CAWI의 무응답률 차이를 의미.

조사방법에 따라 비교하는 등의 추가적인 과정이 필요할 것이다. 한편, 비임금근로자의 경우 CAPI에서 무응답률이 0%였음에도 불구하고 평균소득 무응답률의 유의한 차이는 발견되지 않았다.

임금근로자의 종업원수와 평균소득은 무응답률에 유의한 차이가 있음에도 불구하고 분포와 평균에는 유의한 차이가 없는 것으로 나타났는데, 이는 정확한 수치를 모르는 응답자가 대략적인 수치라도 대답하는 대신 모름/무응답을 선택하기 때문으로 생각된다.

5. 설문유형 실험 분석

CAPI와 CAWI의 경우에는 컴퓨터 화면에서 설문이 어떻게 보이는지에 따라 응답에 영향을 줄 수 있는 요인 중의 하나이다. 예전에 주를 이뤘던 조사도구인 PAPI는 면접원이 주로 설문지를 보면서 응답자에게 질문하고 응답을 기입하는 방식이어서 설문 배열(display)을 크게 고려하지 않아도 되었다. CAPI의 경우도 조사도구만 변경한 경우이기 때문에 비슷할 것이다. 그러나 CAWI의 경우에는 응답자가 컴퓨터 화면에서 직접 설문을 읽고 입력하기 때문에, 응답자가 혼자서도 설문을 잘 응답하거나 쉽게 진행할 수 있도록 화면을 구성하는 것도 중요할 것이다. 그래서 컴퓨터 화면에서 설문 배열 방법의 차이에 따라 응답률에 차이가 있는지 실험해 보았다. 측정 문항은 금융소득, 지난 한 달 소득 및 금융자산으로 선정하였다. 설문유형은 세 가지(A, B, C)를 고려했으며 각각의 설문유형은 다음과 같다. 설문유형 A는 세부 항목 문항을 각각 하나씩 보여주는 것으로, 한 화면에 문항이 1개씩만 보이는 방식이다. 먼저 소득·자산이 존재한다고 응답하면 다음 화면에 금액을 입력하는 것이 나온다. 이 방식이

세부 항목 개수만큼 반복되어 보여준다고 보면 된다. 현재 본 조사의 CAPI 버전에서 이 방식이 사용되고 있다. 설문유형 B는 소득·자산의 유무를 묻지 않고 바로 세부 항목을 전체 화면에 보여주면서 없으면 0원을, 있으면 금액을 입력하도록 구성하였다. 세부 항목의 금액 입력란에 한 항목이라도 비어 있으면 다음 페이지로 넘어가지 않도록 설정하였다. 이 유형은 응답자가 응답과정을 두 번 거치지 않고 한 번에 끝낼 수 있다. 설문유형 C는 우선 세부 항목에 대한 소득·자산의 유무를 한 페이지에서 보여 주어 응답을 받는다. 세부 항목 중에서 소득·자산이 있다고 응답한 목록만 다음 페이지에서 활성화되어 해당 항목에 금액을 입력하도록 구성하였다. 다음은 [그림 7-2]를 통해 세 가지 설문유형을 도식화해 보았다.

〔그림 7-2〕 설문유형 A, B, C에 따라 화면에 보여주는 방식

<table>
<tr><th></th><th>설문유형 A</th><th>설문유형 B</th><th>설문유형 C</th></tr>
<tr><td>(화면1)
존재
유무
문항</td><td><table><tr><th></th><th>존재 유무</th></tr><tr><td>항목 No.</td><td>예/아니오</td></tr></table>화면 한 페이지에서 세부 항목 문항을 하나씩 보여줌.</td><td>- 존재유무 문항이 없고 바로 금액 문항을 보여줌.</td><td><table><tr><th></th><th>존재 유무</th></tr><tr><td>항목 1</td><td>예/아니오</td></tr><tr><td>항목 2</td><td>예/아니오</td></tr><tr><td>항목 3</td><td>예/아니오</td></tr><tr><td>항목 4</td><td>예/아니오</td></tr><tr><td>항목 5</td><td>예/아니오</td></tr></table>- 예를 들어 항목 1, 항목 2와 항목 5를 선택하면 (화면2)는 아래 표와 같이 보여줌.</td></tr>
<tr><td>(화면2)
금액
문항</td><td><table><tr><th></th><th>금액</th></tr><tr><td>항목 No.</td><td>_____만 원</td></tr></table>존재 유무에서 있다고 응답한 경우, 다음 페이지에서 금액 입력하는 문항을 보여줌.
세부 항목 개수만큼 반복함.</td><td><table><tr><th></th><th>금액</th></tr><tr><td>항목 1</td><td>_____만 원</td></tr><tr><td>항목 2</td><td>_____만 원</td></tr><tr><td>항목 3</td><td>_____만 원</td></tr><tr><td>항목 4</td><td>_____만 원</td></tr><tr><td>항목 5</td><td>_____만 원</td></tr></table>해당하는 항목이 없으면 0원, 있으면 금액 입력 화면 한 페이지에서 한 번에 보여줌.</td><td><table><tr><th></th><th>금액</th></tr><tr><td>항목 1</td><td>_____만 원</td></tr><tr><td>항목 2</td><td>_____만 원</td></tr><tr><td>항목 5</td><td>_____만 원</td></tr></table>화면에 선택한 항목만 보여주어서 금액 입력 화면 한 페이지에서 한 번에 보여줌.</td></tr>
</table>

설문유형의 세팅은 2개씩 비교하는 방법을 선택하여 총 세 가지 설문유형 실험을 실시하였다. 금융소득은 설문유형 A와 B를, 지난 한 달 소득은 설문유형 A와 C를, 금융자산은 설문유형 B와 C로 설정한 다음 응답자에게 2개의 설문유형 중 하나를 무작위로 보여주어 응답하도록 하였다.

우선 측정한 3개 문항에 대한 무응답률은 모두 응답하여 0%인 것을 확인하였다. 응답률의 집단 간 차이에 대해 분석을 실시한 결과는 <표 7-45>와 같다. 금융소득은 CAPI와 CAWI에서 모두 설문유형 A와 B 간의 응답률은 통계적으로 유의한 차이가 없음으로 나타났다.

금융소득의 분포를 살펴보았는데 CAWI에서는 설문유형 간 차이가 없는 것으로 나타났으나 CAPI에서는 유의수준 10%에서는 통계적으로 유의한 차이가 있는 것으로 나타났다. 분산의 차이는 CAPI와 CAWI에서 모두 설문유형 간 통계적 차이가 유의하게 없음을 알 수 있다.

다음으로 지난 한 달 소득 역시 CAPI와 CAWI에서 모두 설문유형 A와 C 간의 응답률은 통계적으로 유의한 차이가 없음을 확인하였다.

지난 한 달 소득의 분포도 CAPI와 CAWI에서 모두 설문유형(A : C) 간 통계적으로 유의하게 차이가 없으며, 분산의 분포도 통계적으로 유의한

〈표 7-45〉 금융소득 - 응답률의 집단(A : B) 간 차이 분석

	A		B		diff[1)]	t-value	Pr>|t|
	평균	S.D	평균	S.D			
CAPI	0.3816	0.4890	0.3200	0.4696	0.0616	0.79	0.4313
CAWI	0.4768	0.5011	0.4551	0.4996	0.0217	0.38	0.7044

주 : 1) diff는 CAPI와 CAWI의 평균 차이를 의미.

〈표 7-46〉 금융소득 - 분포 차이 검정

	A		B		diff[1)]	t-value	Pr>|t|	Pr>|Z|[2)]	Levene 검정
	평균 (만 원)	S.D	평균 (만 원)	S.D					
CAPI	433.2	706.9	158.3	164.8	274.8	2.03	0.0510	0.0304	0.1467
CAWI	1,604.6	8,653.0	701.8	1,229.6	902.8	0.88	0.3837	0.3411	0.3117

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.
2) Pr>|Z|는 비모수 검정(nonparametric test)을 통해 계산된 값임.

차이가 없는 것으로 나타났다.

마지막으로, 실험한 금융자산 또한 CAPI와 CAWI에서 모두 설문유형 B와 C 간의 응답률은 통계적으로 유의한 차이가 없음을 알 수 있다.

그러나 금융자산의 분포의 경우 CAWI에서는 비모수 검정에서 유의수준 5%에서 통계적으로 유의한 차이가 있는 것으로 나타났다. CAPI와 CAWI 모두 설문유형(B : C) 간 분산은 통계적으로 유의한 차이가 없는 것으로 나타났다.

설문유형 실험 분석을 실시한 결과 설문유형에 따라 차이가 없어서 특정한 설문유형을 채택할 필요는 없으나, CAWI는 응답자가 직접 컴퓨터 화면을 보면서 설문을 실시하는 것이므로 보기 쉽고 효율적으로 구성

〈표 7-47〉 (지난 한 달) 소득 - 응답률의 집단(A : C) 간 차이 분석

	A		C		diff[1)]	t-value	Pr>|t|
	평균	S.D	평균	S.D			
CAPI	0.8514	0.3582	0.8701	0.3384	-0.0188	-0.33	0.7409
CAWI	0.8243	0.3818	0.8742	0.3327	-0.0499	-1.22	0.2224

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

〈표 7-48〉 (지난 한 달) 소득 - 분포 차이 검정

	A		C		diff[1)]	t-value	Pr>|t|	Pr>|Z|[2)]	Levene 검정
	평균 (만 원)	S.D	평균 (만 원)	S.D					
CAPI	667.2	1,235.6	458.2	454.1	209.0	1.35	0.1806	0.2881	0.2822
CAWI	435.7	449.6	484.0	829.5	-48.2	-0.61	0.5432	0.2245	0.2993

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.
2) Pr>|Z|는 비모수 검정(nonparametric test)을 통해 계산된 값임.

〈표 7-49〉 금융자산 - 응답률의 집단(B : C) 간 차이 분석

	B		C		diff[1)]	t-value	Pr>|t|
	평균	S.D	평균	S.D			
CAPI	0.9474	0.2248	0.9333	0.2511	0.0140	0.36	0.7179
CAWI	0.8408	0.3671	0.9000	0.301	-0.0592	-1.55	0.1224

주 : 1) diff는 CAPI와 CAWI의 평균 차이를 의미.

〈표 7-50〉 금융자산 - 분포 차이 검정

	B		C		diff[1]	t-value	Pr>|t|	Pr>|Z|[2]	Levene 검정
	평균	S.D	평균	S.D					
CAPI	8,024.6만 원	13,391.8	6,770.7만 원	19,960.1	1,253.9	0.44	0.6619	0.1319	0.5613
CAWI	6,993.3만 원	13,240.1	5,851.4만 원	10,365.2	1,142.0	0.78	0.4340	0.0388	0.4743

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.
2) Pr>|Z|는 비모수 검정(nonparametric test)을 통해 계산된 값임.

하도록 고려해야 한다.

6. 직산업 코딩 검정

본항에서는 직산업코드 분포의 조사방법 간 차이를 분석한다. 한국노동패널조사는 응답자들이 종사하는 산업에 관하여 유형 설문에서 “일하는 곳의 주 생산 품목이나 서비스”와 “일하는 곳의 생산품의 원료”를, 직업에 관하여 “주로 하는 일”과 “사용 장비나 도구”, “부서명”을 질문하며, 이를 바탕으로 전문성을 갖춘 코더가 제9차 한국표준산업분류와 제6차 한국표준직업분류에 따른 오픈코드를 지정한다. 한국노동패널조사에서는 직산업코드를 소분류까지 제공하나, 여기서의 분석은 대분류를 이용한다.

직산업코드의 조사방법 간 차이를 분석하기에 앞서 코더 간 불일치율을 살펴볼 필요가 있다. 앞서 설명한 바와 같이, 코더가 수작업으로 오픈코드를 지정하므로 동일한 응답에 대해서 상이한 코드가 배정될 수 있다. 그러나 자료가 충분히 명확하다면 직산업코드의 코더 간 불일치율은 조사방법에 따라 큰 차이를 보이지 않을 것이다. 본 조사에서는 조사방법이 코더의 의사결정에 영향을 미치지 않도록 2명의 코더에게 조사방법에 대한 정보를 알려주지 않았다.

<표 7-51>은 산업코드의 코더 간 불일치율를 보여준다. 임금근로자와 비임금근로자 모두 CAWI에서 CAPI보다 불일치율이 높으며 그 차이가 통계적으로 유의함을 알 수 있다. 앞서 언급한 직산업코드 관련 문항들은 모두 응답자가 주관식으로 입력하는데, 적절한 유도가 없는 경우 답변들이 명확하지 못함을 보여주는 결과이다.

〈표 7-51〉 산업코드의 코더 간 불일치율

	CAPI		CAWI		diff[1)]	t-value	Pr<|t|
	불일치율	S.D.	불일치율	S.D.			
임금근로자	0.0345	0.1831	0.0968	0.2961	-0.0623	-2.75	0.0063
비임금근로자	0.0000	0.0000	0.0870	0.2849	-0.0870	-2.07	0.0442

주 : 1) 산업코드의 코더 간 불일치율의 조사방법에 따른 차이.

〈표 7-52〉 직업코드의 코더 간 불일치율

	CAPI		CAWI		diff[1)]	t-value	Pr<|t|
	불일치율	S.D.	불일치율	S.D.			
임금근로자	0.1862	0.3906	0.1903	0.3932	-0.0041	-0.10	0.9170
비임금근로자	0.0625	0.2459	0.4440	0.2609	-0.1984	-2.52	0.0138

주 : 1) 직업코드의 코더 간 불일치율의 조사방법에 따른 차이.

<표 7-52>는 직업코드의 코더 간 불일치율을 보여준다. 임금근로자는 코더 간 불일치율이 낮고 그 차이가 통계적으로 유의하지 않은 반면, 비임금근로자는 코더 간 불일치율이 높고 그 차이가 통계적으로 유의했다. 임금근로자는 "부서명" 등 보다 명확한 자료가 있는 반면에 비임금근로자는 그렇지 않기 때문으로 추정된다.

가. 임금근로자

<표 7-53>과 <표 7-54>는 각 코더가 지정한 임금근로자의 산업코드 분포의 조사방법 간 차이를 보여준다. 표본의 한계로 인해 정확한 검정은 불가능했으나 산업코드 분포의 조사방법 간 차이는 명백하지 않은 것으로 판단된다. 즉 두 경우 모두 기타서비스업과 광공업이 가장 높은 비중을, 전기·가스 및 수도업과 금융 및 부동산업이 가장 낮은 비중을 차지한 것으로 나타났다(농림어업 제외). 한편, CAWI에서 CAPI에 비해 광공업의 비중은 높은 반면에 공공서비스업과 기타서비스업의 비중은 낮은 것으로 나타났다.

<표 7-55>와 <표 7-56>은 각 코더가 지정한 임금근로자의 직업코드

분포의 조사방법 간 차이를 보여준다. 산업코드와 마찬가지로, 직업코드 분포 역시 조사방법 간 차이가 명백하지 않은 것으로 보인다. 즉 전문관

〈표 7-53〉 임금근로자의 산업코드 분포(코더 1)

	CAPI		CAWI		diff[1)]
	빈도	비중	빈도	비중	
농림어업	0	0.0	0	0.0	0.0
광공업	30	20.7	76	24.9	-4.2
건설업	10	6.9	23	7.5	-0.6
전기·가스 및 수도업	1	0.7	2	0.7	0.0
도소매 및 음식숙박업	9	6.2	27	8.9	-2.7
운수통신업	17	11.7	30	9.8	1.9
금융 및 부동산업	8	5.5	22	7.2	-1.7
공공서비스업	12	8.3	14	4.6	3.7
기타서비스업	58	40.0	111	36.4	3.6
전체	145	100.0	305	100.0	

주 : 1) 산업코드 비중의 조사방법에 따른 차이.

〈표 7-54〉 임금근로자의 산업코드 분포(코더 2)

	CAPI		CAWI		diff[1)]
	빈도	비중	빈도	비중	
농림어업	0	0.0	1	0.3	-0.3
광공업	30	20.7	73	23.9	-3.2
건설업	11	7.6	26	8.5	-0.9
전기·가스 및 수도업	1	0.7	2	0.7	0.0
도소매 및 음식숙박업	9	6.2	21	6.9	-0.7
운수통신업	18	12.4	36	11.8	0.6
금융 및 부동산업	8	5.5	24	7.9	-2.4
공공서비스업	12	8.3	12	3.9	4.4
기타서비스업	56	38.6	110	36.1	2.5
전체	145	100.0	305	100.0	

주 : 1) 산업코드 비중의 조사방법에 따른 차이.

〈표 7-55〉 임금근로자의 직업코드 분포(코더 1)

	CAPI		CAWI		diff[1)]
	빈도	비중	빈도	비중	
전문관리직	59	40.7	126	40.9	-0.2
사무직	50	34.5	124	40.3	-5.8
서비스직	16	11.0	17	5.5	5.5
농림어업직	0	0.0	0	0.0	0.0
생산직	20	13.8	41	13.3	0.5
전체	145	100.0	308	100.0	

주 : 1) 직업코드 비중의 조사방법에 따른 차이.

〈표 7-56〉 임금근로자의 직업코드 분포(코더 2)

	CAPI		CAWI		diff[1)]
	빈도	비중	빈도	비중	
전문관리직	63	43.5	133	43.2	0.3
사무직	49	33.8	116	37.7	-3.9
서비스직	13	9.0	21	6.8	2.2
농림어업직	0	0.0	0	0.0	0.0
생산직	20	13.8	38	12.3	1.5
전체	145	100.0	308	100.0	

주 : 1) 직업코드 비중의 조사방법에 따른 차이.

리직과 사무직이 높은 비중을 차지하고 나머지 직종들의 비중은 비교적 낮은 것으로 나타났다. 그러나 두 경우 모두 CAWI에서 CAPI보다 사무직 비중은 높은 반면에 서비스직 비중은 낮았으며, 그 차이도 비교적 큰 것으로 보인다.

나. 비임금근로자

<표 7-57>과 <표 7-58>은 각 코더가 지정한 비임금근로자의 산업코드 분포의 조사방법 간 차이를 보여준다. 임금근로자와 달리, 비임금근로자의 산업코드 분포는 조사방법에 따라 유의한 차이가 있는 것으로 보인

다. 기타서비스업과 도소매 및 음식숙박업의 비중이 가장 큰 것은 동일하나 그 차이가 크고, 그 외 직업코드들의 분포도 상이하다. 단, 이는 비임

〈표 7-57〉 비임금근로자의 산업코드 분포(코더 1)

	CAPI		CAWI		diff[1)]
	빈도	비중	빈도	비중	
농림어업	2	6.5	0	0.0	6.5
광공업	4	12.9	5	11.1	1.8
건설업	1	3.2	0	0.0	3.2
전기·가스 및 수도업	0	0.0	0	0.0	0.0
도소매 및 음식숙박업	9	29.0	16	21.5	7.5
운수통신업	0	0.0	5	11.1	-11.1
금융 및 부동산업	3	9.7	6	13.3	-3.6
공공서비스업	0	0.0	0	0.0	0.0
기타서비스업	12	38.7	13	28.9	9.8
전체	31	100.0	45	100.0	

주 : 1) 산업코드 비중의 조사방법에 따른 차이.

〈표 7-58〉 비임금근로자의 산업코드 분포(코더 2)

	CAPI		CAWI		diff[1)]
	빈도	비중	빈도	비중	
농림어업	2	6.5	0	0.0	6.5
광공업	4	12.9	8	17.8	-4.9
건설업	1	3.2	0	0.0	-3.2
전기·가스 및 수도업	0	0.0	0	0.0	0.0
도소매 및 음식숙박업	9	29.0	13	28.9	0.1
운수통신업	0	0.0	5	11.1	-11.1
금융 및 부동산업	3	9.7	5	11.1	-1.4
공공서비스업	0	0.0	0	0.0	0.0
기타서비스업	12	38.7	14	31.1	7.6
전체	31	100.0	45	100.0	

주 : 1) 산업코드 비중의 조사방법에 따른 차이.

금근로자의 표본이 충분하지 않아 약간의 빈도 차이로도 비중 차이가 클 수 있기 때문일 수 있다.

<표 7-59>와 <표 7-60>은 각 코더가 지정한 비임금근로자의 직업코드 분포의 조사방법 간 차이를 보여준다. 산업코드와 마찬가지로, 직업코드 분포 역시 조사방법 간 차이가 유의한 것으로 보인다. 코더 1의 경우 CAPI에서는 서비스직 비중이 가장 높은 반면에 CAWI에서는 전문관리직의 비중이 가장 높으며, 그 외 직업코드들의 분포도 상이하다. 한편, 코더 2의 경우 전반적인 분포는 유사하나 각 비중의 조사방법 간 차이가 큰 것으로 나타났다. 단, 앞서 언급한 바와 같이, 비중의 차이가 큰 것으로 나타난 것은 표본수의 문제일 수 있다.

〈표 7-59〉 비임금근로자의 직업코드 분포(코더 1)

	CAPI		CAWI		diff[1)]
	빈도	비중	빈도	비중	
전문관리직	10	32.3	20	44.4	-12.1
사무직	0	0.0	3	6.7	-6.7
서비스직	15	48.4	17	37.8	10.6
농림어업직	2	6.5	0	0.0	-6.5
생산직	4	12.9	5	11.1	1.8
전체	31	100.0	45	100.0	

주 : 1) 직업코드 비중의 조사방법에 따른 차이.

〈표 7-60〉 비임금근로자의 직업코드 분포(코더 2)

	CAPI		CAWI		diff[1)]
	빈도	비중	빈도	비중	
전문관리직	10	32.3	13	28.9	3.4
사무직	1	3.2	8	17.8	-14.6
서비스직	14	45.2	15	33.3	11.9
농림어업직	2	6.5	0	0.0	6.5
생산직	4	12.9	9	20.0	-7.1
전체	31	100.0	45	100.0	

주 : 1) 직업코드 비중의 조사방법에 따른 차이.

- 주된 일자리에서는 세 가지 항목으로 자세하게 묻고,
- 14세경 부모님 일자리, 회고적 일자리는 간략하게 묻는데, 여기서 불성실한 응답이 많이 보임.(특히 CAWI에서)_→분석 추가할 예정

다. 소 결

본항에서는 CAPI와 CAWI의 직산업코드의 분포가 조사방법 간 차이가 유의한지 분석하였다. 먼저 직산업코드의 코더 간 불일치율의 조사방법 간 차이를 살펴본 결과, 코더 간 불일치율은 CAWI에서 CAPI보다 높은 것으로 나타났다. 이를 통해 CAWI를 통해 조사된 자료가 코더들에게 명확한 정보를 주지 못하고 있음을 추론할 수 있다. 또한 전반적으로, 임금근로자의 직산업코드 분포는 조사방법 간 차이가 상대적으로 작은 반면에 비임금근로자는 큰 것으로 나타났다. 이러한 결과는 비임금근로자의 표본수가 적었기 때문일 수 있으나, 코더 간 불일치율도 비임금근로자가 임금근로자보다 큰 것으로 나타나 CAWI를 통해 조사된 자료가 코더들에게 충분한 정보를 주지 못했을 가능성을 추론할 수 있다. 이는 향후 CAPI에서 CAWI로 이행함에 있어 오픈코드 지정을 위한 주관식 답변에 대해 적절한 가이드라인 제시가 중요함을 보여주는 결과이다.

제4절 소 결

본 연구는 노동패널조사의 CAWI조사방법의 도입 가능성 여부를 타진하기 위한 것이다. CAPI 151가구, CAWI 307가구의 비교분석을 통해 다음과 같은 결과를 얻었다.

첫째, CAWI는 CAPI에 비해 조사시작은 용이하지만, 마지막까지 조사를 완료하는 비율은 현저히 낮았다. 또한 가구에 속한 15세 이상 가구원 전수를 조사해야 함에도 불구하고, 응답을 거부하는 가구원이 많아 개인의 응답률이 저조하였고, 대리응답도 더 많은 것으로 나타났다. 성별이나

학력에 따라서는 차이를 보이지 않지만, CAWI 응답자의 연령이 더 낮은 것으로 조사되었다.

둘째, 항목응답률에서는 가구소득, 금융자산, 지난 한 달 근로소득, 월평균 저축액, 금주 등의 변수는 CAPI의 응답률이 통계적으로 유의하게 더 높았다.

셋째, 분포의 차이를 살펴보면, 가구소득변수(근로소득액, 금융소득액, 부동산소득액, 이전소득액, 기타소득액)는 근로소득액만 통계적으로 유의한 차이를 보였다. 그 외 가구 변수에서는 금융자산과 지난 한 달 소득, 월평균 저축액은 차이가 있는 것으로 조사되었다.

개인응답 설문에서는 만족도 문항이 대부분 통계적으로 유의한 차이를 보였는데, 전반적인 일자리(직장)만족도, 전반적 일(직무)만족도, 전반적 생활만족도에서는 CAWI 평균값이 더 높았고, 세부 항목에서도 통계적으로 유의한 차이를 발견되었다. 직업훈련 경험여부와 사회보험 수급여부도 분포의 차이를 보였다. 흡연과 금연시도, 금주시도는 분포에서도 차이를 보였다. 일자리 관련 변수 중 고용형태와 평균근로시간은 통계적으로 유의한 차이를 보이지 않았지만, 통계적 기준의 비정규직 여부에서 차이가 나타났다. 이는 '근로기간이 정함이 있느냐'는 질문에서 응답의 차이를 보였기 때문으로 파악되는데, 계약기간이 10년 이상으로 응답한 경우가 대부분으로 정년과 근로기간 정함의 개념을 오해하여 응답한 것으로 생각된다.

넷째, 항목 무응답률은 가구변수, 교육, 혼인상태, 일자리시작시기 등의 변수는 조사방법에 따라 차이가 없는 것으로 조사되었다. 그러나 세전 근로소득, 세후 근로소득, 종업원수는 CAWI의 무응답이 더 높은 것으로 나타났다.

다섯째, CAWI 방법 내에서 설문유형에 따른 차이는 금융소득, 지난 한 달 소득, 금융자산 세 가지 항목에 대해 모두 통계적으로 유의한 차이가 없어서 특정한 설문유형을 채택할 필요는 없는 것으로 나타났다. 그러나 응답자의 편리를 위해 보기 쉽고 효율적으로 구성하도록 고려해야 한다.

여섯째, 직산업 코딩은 2명의 코더가 각자 작성하여 그 차이를 검정하였다. 대분류 기준으로 산업은 CAPI보다 CAWI에서 임금근로자, 비임금

근로자에서 모두 불일치가 높게 발생하였고, 직업은 비임금근로자에 한해 차이가 발생하였다.

이상의 분석결과는 조사성공률 측면에서 CAWI가 CAPI보다 효과적이지 않은 것으로 생각된다. 항목응답률, 분포의 차이, 항목 무응답률, 직산업 코딩 등의 검정을 통해서는 일부 항목들에서는 조사 시행 전에 설문문항의 보완이나 후속 전화조사의 보완이 필요한 것으로 보인다.

패널조사의 가장 중요한 점은 원가구를 유지하여 패널의 성격을 유지하는 것이다. 그러므로 이사추적불가 가구가 늘어나고, 청년층 응답자가 계속 줄어드는 것은 반드시 해결해야 할 문제이다. 따라서 CAWI의 전면적인 확대를 목표로 두기보다는, 비성공 가구들에 대한 보완적인 조사방법 중의 하나로서의 접근은 계속적으로 시도해 보아야 할 것이다.

참고문헌

강석훈(2003), 『KLIPS의 가중치 부여방안 연구』, 한국노동패널연구 2003-04, 한국노동연구원.

강성호(2014), 「연금세제 개편에 따른 개인연금의 체감보험료 변화효과」, 2014년 재정패널 학술대회.

고제이 · 권혁진 · 신우진 · 류재린 · 하솔잎 · 조남운(2016), 『미시모의실험 기반 중장기 사회 재정 영향 평가 모형 개발 - 노후 소득 보장 정책을 중심으로』, 한국보건사회연구원.

국회예산정책처(2017), 『2017 조세의 이해와 쟁점 : ②소득세』, 국회예산정책처.

권혁진(2007), 「복지혜택 지급방식에 따른 조세, 복지제도의 불평등완화 효과와 재정부담에 대한 시뮬레이션 분석」, 『사회보장연구』 23(2), pp.29~59.

_____(2016), 「노동패널자료 소득세액 추정」, 홍민기 외 5인, 『패널자료 품질개선 연구(VI)』 제5장, 한국노동연구원.

권혁진 · 류재린(2015), 「공적연금의 최저생계 보장 효과에 대한 장기 전망」, 「응용통계연구」 28(4), pp.741~762.

권혁진 · 신우진 · 한재명 · 정동진 · 류재린 · 박충렬 · 김균(2013), 「2007년 공적 연금 개혁의 적정성 전망과 평가 - dynamic micro-simulation 을 통한 장기 분석-」, 『사회보장연구』 29(4), pp.1~54.

권혁진 · 한정림(2009), 『국민연금의 미시모의실험 모형 연구 및 개발 - KMAP ver.1.0』 국민연금연구원.

기획재정부(2016), 『2016 조세개요』, 기획재정부.

김기민(2009), 「CAPI 조사리뷰 프로그램 소개」, 『노동리뷰』 2009년 3월호(통권 제51호), 한국노동연구원, pp.56~62.

김대환(2017), 「사적연금 세액공제 확대에 따른 소득계층별 영향과 정책

적 시사점」, 2017년 재정패널 학술대회.
김명규(2017), 「근로소득공제제도의 결혼·육아에 대한 효과」, 2017년 재정패널 학술대회.
김민성 · 오시진(2010), 「평률세제 개편안의 정치적 실현 가능성과 조세부담의 재분배」, 『재정학연구』 3(3), pp.39~71.
김병권 · 우석진 · 안종길 · 빈기범(2013), 「세제혜택이 가계의 연금저축 행태에 미치는 인과적 효과」, 『사회보장연구』 29(3), pp.53~79.
김성태 · 임병인(2013), 「재정패널자료를 이용한 특별공제와 세액공제의 형평성 효과 비교 분석」, 2013년 재정패널 학술대회.
김영원 · 김재광 · 이기재 · 조유미(2005), 「한국노동패널 표본의 대표성과 가중치 보정방법」, 『제6회 한국노동패널 학술대회 논문집』, pp.323~339.
김은정 · 이혜숙(2016), 「영유아 보육(돌봄) 지원 정책 평가와 정책 과제-보육료 지원 정책을 중심으로」, 한국보건사회연구원.
김재광(2003), 『한국노동패널 오차분석 및 표본추가』, 한국노동패널연구 2003-02, 한국노동연구원.
_____(2009), 『한국노동패조사 표본 추가 연구용역 최종보고서』, 한국노동연구원.
김재호(2015), 「소득공제혜택이 개인연금 보험료에 미치는 효과」, 2015년 재정패널 학술대회.
남재량(2017), 「근로장려세제의 노동시장 성과 및 최저임금제와의 관계 연구」, 제19회 노동패널 학술대회.
민연경 · 장한나(2015), 「보육정책의 이슈네트워크 분석 : 무상보육정책과 양육수당정책을 중심으로」, 『한국행정연구』 24(2), pp.97~138.
박명호(2016), 「기부금 조세정책이 개인의 자발적 기부행위에 미치는 영향」, 2016년 재정패널 학술대회.
박민규 · 김영원 · 최형아(2011), 「13차년도 한국노동패널 가중치 연구」, 한국고용정보원.
박민규 · 김원 · 변종석 · 송종호 · 임찬수 · 김사라(2013), 『한국노동패널 가중치 연구』, 한국노동연구원.

박종선・정세은(2017), 「2008년 이후 소득세제 개편의 소득분배 및 세부담 효과」, 2017년 재정패널 학술대회.

보건복지부・한국보건사회연구원(2017), 『통계로 보는 사회보장 2016』.

서수경(2012), 「어린이집 특별활동 및 필요경비 가이드라인 구축」, 서울시여성가족재단.

선우희연・전병욱・정운오(2015), 「한계조세부담 및 금융소득종합과세의 강화가 가계의 금융자산 포트폴리오에 미치는 영향」, 2015년 재정패널 학술대회.

성명재・송헌재・전병목(2010), 『조세・재정모의실험 모형 : KIPFSIM10 모형의 구축』, 한국조세연구원.

성명재・전병목・전병힐(2008), 『조세・재정모의실험 모형 : KIPFSIM08 모형의 구축』, 한국조세연구원.

송헌재(2013), 「재정패널의 소득증빙자료를 활용한 근로소득자들의 기부금 가격탄력성 추정」, 2013년 재정패널 학술대회.

송헌재・성명재(2011), 「신용카드 사용액 소득공제 제도의 효과 분석」, 2011년 재정패널 학술대회.

송헌재・신우리(2017), 「소득공제의 세액공제 전환이 노동공급에 미친 영향 분석」, 2017년 재정패널 학술대회.

신현구・이혜정(2006), 「컴퓨터를 이용한 대인면접이 조사 자료의 질 개선에 미치는 영향 : 고령화연구패널조사의 Blaise 활용사례를 중심으로」, 『조사연구』 7(2), pp.71~95.

심욱기(2006), 「2002년 세율인하 정책을 이용한 노동공급의 임금탄력성 분석」, 『2006년 노동패널 학술대회』.

안현미・박애리・남은이(2013), 「무상보육(보육료 전액지원)과 보육서비스 이용 실태 - 서울시 영유아 부모를 중심으로」, 서울시여성가족재단.

이계오・임찬수(2012), 「한국노동패널 가중치 산출방안 - 원패널과 추가패널」, 『한국노동패널 학술대회 논문집』.

전병욱(2016), 「소득공제 대상인 특별공제 항목의 세액공제 전환에 따른 지출행태의 변화 분석」, 2016년 재정패널 학술대회.

전병욱(2017),「가구별 종교가 절세효과 및 세액공제 전환에 따른 기부의 사결정에 미치는 영향의 분석」, 2017년 재정패널 학술대회.

전병욱・선우희연・정운오(2015),「절세효과가 기부금 지출 의사결정에 미치는 영향: 공제한도초과 기부금에 미치는 한계세율의 영향을 중심으로」, 2015년 재정패널 학술대회.

전승훈・강성호・임병인(2008),「개인연금 불입액의 가처분소득 탄력성과 세제혜택」, 2008년 노동패널 학술대회.

정원석(2017),「소득수준에 따른 세액공제율 차등적용 효과 분석」, 2017년 재정패널 학술대회.

정원석・강성호(2013),「개인연금 소득공제 제도의 세액공제로의 변화 시 소득계층별 정책효과분석」, 2013년 재정패널 학술대회.

_____(2015),「세제혜택이 사적연금을 증가시키는가」, 2015년 재정패널 학술대회.

_____(2016),「연금저축 세제혜택 세액공제 전환에 따른 소득계층별 연금저축 납입효과 분석」, 2016년 재정패널 학술대회.

최보람(2015),「근로소득세에 관한 연구 - 2015년 연말정산 파동을 중심으로」, 한국국제회계학회,『국제회계연구』63, pp.227~252.

최성은(2016),「보육 및 유아교육 비용과 보육료 상한 규제」,『재정포럼』235, 현안분석2, pp.23~35.

최효미・김길숙・이동하・임준범(2016).「영유아 교육・보육비용 추정연구(Ⅳ)」, 육아정책연구소.

최효미・성재민・이상호・배기준(2012),「한국노동패널조사(KLIPS)의 CAPI 이행효과: 실험설계 및 차수간 정합성을 중심으로」,『조사연구』13(3), pp.105~135.

최효미・이상호・성재민・김기민・배기준(2012),『패널자료의 품질개선을 위한 조사기법 연구 - 컴퓨터를 이용한 조사(CAPI) 방법』, 한국노동연구원.

한국리서치(2016), 한국리서치 응답자(MS패널) 소개.

홍민기・강신욱・성재민・이상호・장지연・김재광・이지은(2015),「노동패널자료연구(1): 패널자료 비교연구를 중심으로」, 한국노동연구원.

홍민철・문상호・이명석(2016),「근로장려세제 효과 분석 : 경제활동참여, 근로시간 및 개인별 빈곤을 중심으로」, 한국정책분석평가학회,『정책분석평가학회보』26(2), pp.1~27.

Agur, Maya, Curtis Jessop, Sarah Kunz, Violetta Parutis, and Martin Wood(2014), "Understanding Society Innovation Panel Wave 6, Technical Report," The Institute for Social and Economic Research, University of Essex.

Al Baghal, Tarek, Peter Lynn(2015), "Using Motivational Statements in Web-Instrument Design to Reduce Item-Missing Rates in a Mixed-Mode Context," *Public Opinion Quarterly* 79(2), pp.568~579.

Atkinson, A. B., F. Bourguignon, and P. A. Chiappori(1988), "What Do We Learn about Tax Reforms from International Comparisons? France and Britain," *European Economic Review* 32(2-3), pp.343~352.

Bargain, O. and T. Callan(2010), "Analysing the Effects of Tax-Benefit Reforms on Income Distribution: A decomposition approach," *Journal of Economic Inequality* 8(1), pp.1~21.

Bargain, O., A. Decoster, M. Dolls, D. Neumann, A. Peichl, and S. Siegloch(2012a), "Welfare, Labor Supply and Heterogeneous Preferences: Evidence for Europe and the US," *Social Choice and Welfare,* Forthcoming.

Bargain, O., M. Dolls, C. Fuest, D. Neumann, A. Peichl, N. Pestel, and S. Siegloch(2012b), "Fiscal Union in Europe? Redistributive and Stabilising Effects of an EU Tax-Benefit System," IZA Discussion Paper 6585.

Bargain, O., M. Dolls, D. Neumann, A. Peichl, and S. Siegloch(2011), "Tax-Benefit Systems in Europe and the US: Between equity and efficiency," IZA Discussion Paper 5440.

Barrios, S., S. Fatica, D. Martinez, and G. Mourre(2015), "The Fiscal Effects of Work-Related Tax Expenditures in Europe," European Economy Economic Papers 545, Directorate General Economic and Financial Affairs, European Commission.

Brewer, M., J. Browne, C. Emmerson, A. Goodman, A. Muriel, G. Tetlow(2007), "Pensioner Poverty over the Next Decade : What role for tax and benefit reform?", IFS Commentary No.103, The Institute for Fiscal Studies.

Callan, T. and H. Sutherland(1997), "The Impact of Comparable Policies in European Countries : Microsimulation approaches," *European Economic Review* 41, pp.627~633.

Canto, O., M. Adiego, L. Ayala, H. Levy, and M. Paniagua(2012), "Going Regional. The effectiveness of different tax-benefit policies in combating child poverty in Spain," EUROMOD Working Paper Series EM2/12.

CBO(2012), "Effective Marginal Tax Rates for Low- and Moderate-Income Workers", CBO Report, 2012-11-15. (https://www.cbo.gov/publication/43709)

de Leeuw, Edith D. and Joop J. Hox(2011), "Internet Surveys as Part of a Mixed-Mode Design," in Marcel Das, Peter Ester, and Lars Kaczmirek(eds.), *Social and Behavioral Research and the Internet,* Taylor and Francis Group.

_____(2015), "Survey Mode and Mode Effects," in Uwe Engel, Ben Jann, Peter Lynn, Annette Scherpenzeel, and Ptrick Sturgis(eds.), *Improving Survey Methods - Lessons from Recent Research,* Routledge : Taylor & Francis

Decoster, A., J. Loughrey, C. O'Donoghue, and D. Verwerft(2010), "How Regressive Are Indirect Taxes? A microsimulation analysis for five European countries," *Journal of Policy Analysis and Management* 29(2), pp.326~350.

Ericson, P, L.R Flood, R. Wahlberg(2009), "SWEtaxben : A Swedish tax/benefit micro simulation model and an evaluation of a Swedish tax reform," IZA DP No.4106.

EUROMOD(2017), "EUROMOD Modelling Conventions," EUROMOD Technical Note EMTN 1.0, Colchester: Institute for Social and Economic Research, University of Essex.

Figari, F., H. Levy, and H. Sutherland(2007), "Using the EU-SILC for Policy Simulation : Prospects, some limitations and some suggestions", in Comparative EU-Statistics on Income and Living Conditions : Issues and Challenges, Eurostat Methodologies and Working Papers, European Communities.

Figari, F., A. Paulus, H. Sutherland, P. Tsakloglou, G. Verbist and F. Zantomio(2012b), "Taxing the Benefit of Homeownership. Distributional effects of including imputed rent in taxable income," IZA Discussion Paper No.6493.

Gerstorf, Sandra, Jürgen Schupp(eds.)(2014), "SOEP Wave Report," DIW Berlin, Germany.

Hanson, Tim, Peter Matthews, and Alice McGee(2016), "Understanding Society Innovation Panel Wave 8, Technical Report."

Immervoll, H.(2004), "Average and Marginal Effective Tax Rates Facing Workers in the EU. A micro-level analysis of levels, distributions and driving factors," OECD Social, Employment and Migration Working Paper No.19, OECD: Paris.

Immervoll, H. and C. O'Donoghue(2004), "What Difference Does a Job Make? The Income Consequences of Joblessness in Europe," in D. Gallie(ed.), *Resisting Marginalisation : Unemployment experience and social policy in the European Union,* Oxford: Oxford University Press.

Immervoll, H., H. J. Kleven, C. T. Kreiner, and E. Saez(2007), "Welfare Reform in European Countries : A microsimulation analysis," *The*

Economic Journal 117(516), pp.1~44.

Immervoll, H., H. J. Kleven, C. T. Kreiner, and N. Verdelin(2011), "Optimal Tax and Transfer Programs for Couples with Extensive Labor Supply Responses," *Journal of Public Economics* 95(11-12), pp.1485~1500.

Immervoll, H., H. Levy, C. Lietz, D. Mantovani, C. O'Donoghue, H. Sutherland and G. Verbist(2006), "Household Incomes and Redistribution in the European Union : Quantifying the equalizing properties of taxes and benefits," in D. B. Papadimitriou(ed.), *The Distributional Effects of Government Spending and Taxation,* Palgrave Macmillan.

Islam, N. and L. Flood(2015), "A Tax Benefit Model for Policy Evaluation in Luxembourg: LuxTaxBen," IZA DP No.9152.

Jäckle, Annette, Peter Lynn, and Jon Burton(2013), "Going Online with a Face-to-Face Household Panel: Initial Results from an Experiment on the Understanding Society Innovation Panel," Institute for Social and Economic Research, University of Essex.

Jara, X. and A. Tumino(2013), "Tax-Benefit Systems, Income Distribution and Work Incentives in the European Union," *The International Journal of Microsimulation* 6(1), pp.27~62.

Klevmarken, N.A.(1991), "MICROHUS: A Microsimulation Model for the Swedish Household Sector," A progress report, International Geographical Union Microsimulation Symposium, Stockholm, Sweden, pp.2~4.

Klevmarken, N.A., P. Olovsson(1996), "Direct and Behavioural Effects of Income Tax Changes-Simulations with the Swedish Model Microhus," in A. Harding(ed.), *Microsimulation and Public Policy*, Contributions to Economic Analysis No 232. Elsevier Science B. V., Amsterdam, the Netherlands.

Levy, H., C. Lietz, and H. Sutherland(2007), "Swapping Policies :

Alternative tax-benefit strategies to support children in Austria, Spain and the UK," *Journal of Social Policy* 36(4), pp.625~647.

Levy, H., L. Morawski, and M. Myck(2009), "Alternative Tax-benefit Strategies to Support Children in Poland," in O. Lelkes and H. Sutherland(eds.), *An Enlarged Role for Tax Benefit Models: Assessing policies in the enlarged European Union,* Farnham: Asghate.

Li, J. and C. O'Donoghue(2013), "A Survey of Dynamic Microsimulation Models: Uses, model structure and methodology," *International Journal of Microsimulation* 6(2), pp.3~55.

Lohr, S. L. and J. N. K. Rao(2000), "Inference from Dual Frame Surveys," *Journal of the American Statistical Association* 95(449), pp.271~280.

Lynn, P.(2009), "Methods for Longitudinal Surveys," in P. Lynn(ed.), *Methodology of Longitudinal Surveys,* Chichester: John Wiley and Sons.

Maestri, V.(2012), "Economic Well-being and Distributional Effects of Housing-related Policies in 3 European Countries," EUROMOD Working Paper Series EM10/12.

McGonagle, Katherine A., Vicki A. Freedman, Jamie Griffin, Mary Dascola(2017), "Web Development in the PSID: Translation & Testing of a Web Version of the 2015 PSID Telephone Instrument," Technical Series Paper #17-02, PSID.

Merz, J., H. Stolze, M. Zwick(2002), "Professions, entrepreneurs, employees and the new German tax (cut) reform 2000-A MICSIM microsimulation analysis of distributional impacts," FFB Discussion Paper No.34.

Ofstedal, Mary Beth, Mick P. Couper, and Colleen McClain(2016), "Transitioning a Panel Survey to Web: Experiences from the health and retirement study," Working Paper.

O'Muircheartaigh, C. and S. Pedlow(2002), "Combining Samples vs. Cumulating Cases : A comparison of two weighting strategies in NLSY97." *American Statistical Association Proceedings of the Joint Statistical Meetings,* pp.2557~2562.

Paulus A. and A. Peichl(2009), "Effects of Flat Tax Reforms in Western Europe," *Journal of Policy Modeling* 31(5), pp.620~636.

Peichl, A., H. Schneider, S. Siegloch(2010), "Documentation IZAΨMOD : The IZA Policy SImulation MODel," IZA DP No. 4865.

PSID Main Interview User Manual: Release(2015), Institute for Social Research, University of Michigan, June, 2015.

Smyth, Jolene D. and Jennie E. Pearson(2011), "Internet Survey Methods: A review of strengths, weaknesses, and innovations," in Marcel Das, Peter Ester, and Lars Kaczmirek(eds.), *Social and Behavioral Research and the Internet*, Taylor and Francis Group.

Sutherland, H.(1997), "The EUROMOD Preparatory Study : A summary report," Microsimulation Unit Discussion Paper, MU9705, University of Cambridge.

_____(2016), "EUROMOD : Present, past and future," EUROMOD 20th Anniversary Conference. (download at https : //www.iser.essex.ac.uk/files/euromod/euromod20/pdf/SUTHERLAND.pdf)

Sutherland, H. and F. Figari(2013), "EUROMOD : The European Union tax-benefit microsimulation model," *Internatonal Journal of Microsimulation* 6(1), pp.4~26.

Taylor, Marcia Freed(ed.), with John Brice, Nick Buck, and Elaine Prentice-Lane(2010), "British Household Panel Survey User Manual Volume A : Introduction, Technical Report and Appendices," Colchester : University of Essex.

Watson, N.(2012), "Longitudinal and Cross-Sectional Weighting Methodology for the HILDA Survey," HILDA Technical Paper Series 2(12).

_____(2014), “Evaluation of Weighting Methods to Integrate a Top-up Sample with an Ongoing Longitudinal Sample,” *Survey Research Methods* 8(3), pp.195~208.

Zaidi, A., K. Rake(2001), “Dynamic Microsimulation Models : A review and some lessons for SAGE,” The London School of Economics, Simulating Social Policy in an Ageing Society(SAGE), Discussion Paper.

Zaidi, A., A. Scott(2001), Base dataset for the SAGE model, Sage Technical Note.

http://about.paneldata.org/soep/dtc/sample.html

◈ 執筆陣

• 김유빈(한국노동연구원 연구위원)
• 박민규(고려대학교 교수)
• 변종석(한신대학교 교수)
• 최효미(육아정책연구소 부연구위원)
• 권혁진(경남과학기술대학교 부교수)
• 김우영(공주대학교 교수)
• 성재민(한국노동연구원 연구위원)
• 이지은(한국노동연구원 전문위원)
• 신선옥(한국노동연구원 책임연구원)
• 이혜정(한국노동연구원 전문위원)
• 김기홍(한국노동연구원 연구원)

패널자료 품질개선 연구(Ⅶ)

▪ 발행연월일	2017년 12월 26일 인쇄 2017년 12월 29일 발행
▪ 발 행 인	김 승 택 원장직무대행
▪ 발 행 처	**한국노동연구원** 30147 세종특별자치시 시청대로 370 세종국책연구단지 경제정책동 ☎ 대표 (044) 287-6080 Fax (044) 287-6089
▪ 조판・인쇄	거목정보산업(주) (044) 863-6566
▪ 등 록 일 자	1988년 9월 13일
▪ 등 록 번 호	제13-155호

ⓒ 한국노동연구원 2017

정가 17,000원

ISBN 979-11-260-0172-9