정책 21-04

OPEN 공공누리 출처표시 상업용금지 변경금지 공공저작물 자유이용허락

장애인삶 패널조사 패널이탈 특성 및 종단가중치(3차년도) 연구

2021. 10.

한국장애인개발원
Korea Disabled people's Development Institute

본 보고서의 연구내용은 연구진의 의견으로 한국장애인개발원의 공식적인 의견과는 다를 수 있습니다.

연 구 진

연 구 책 임 : 김현지 (한국장애인개발원 정책연구부 조사패널팀 팀장)

공동연구원 : 김태용 (한국장애인개발원 정책연구부 조사패널팀 선임연구원)
이혜수 (한국장애인개발원 정책연구부 조사패널팀 연구원)

자문위원

변 종 석 (한신대학교 응용통계학과 교수)
박 민 규 (고려대학교 통계학과 교수)
임 찬 수 (통계청 통계개발원 사무관)

발간사

한국장애인개발원은 2018년부터 장애인삶 패널조사(승인번호: 제438001호)를 실시하고 있습니다. 저희 패널조사는 2015년부터 2017년 사이에 장애등록을 마친 사람을 모집단으로, 장애등록 이후 삶의 변화를 파악하고, 관련 정책의 수립·지원에 필요한 기초자료 마련을 목적으로 하고 있습니다. 이에 2018년 6,121명의 장애인 패널을 구축하였습니다.

3차 조사는 장애인 5,259명과 그들의 가구원이 조사에 참여함으로서, 장애등록 이후 장애인의 삶의 변화를 통합적으로 분석할 수 있는 종단자료의 형태를 갖추게 되었습니다. 이에 본 패널조사의 조사 결과로 생산되는 데이터를 활용한 연차 비교 및 분석에 필요한 종단가중치(3차년도) 작성 방안 연구를 수행하게 되었습니다. 이를 통해 시간 경과에 따른 개인의 변화 등을 파악할 수 있을 것이라 기대합니다. 뿐만 아니라 3차 조사의 응답결과를 기준으로 응답패턴, 응답 패널과 무응답 패널간 이탈 특성 분석을 실시하여 데이터 품질 관리를 위한 분석 기반을 마련하였습니다.

본 조사가 장애인의 장애등록 이후 장애수용 과정과 사회관계 속에서 겪는 변화를 다양한 측면에서 이해하는 디딤돌이 되었으면 하는 바람입니다. 장애인삶 패널조사에 여러분들의 많은 관심과 조언을 부탁드립니다. 끝으로 본 연구의 진행에 아낌없는 조언과 자문을 주신 학계 및 관계자 분들께도 깊은 감사를 드립니다.

2021년 10월

한국장애인개발원 원장 최 경 숙

한국장애인개발원
www.koddi.or.kr

목 차

연구요약 ····· ix

Ⅰ. 서론 ····· 1

1. 개요 ····· 3
2. 연구필요성 및 목적 ····· 5
3. 연구내용 및 방법 ····· 7
   1) 연구내용 ····· 7
      (1) 개요 ····· 7
      (2) 연구의 주요 내용 ····· 7
   2) 연구 수행 방법 ····· 8
      (1) 기본프레임 ····· 8
      (2) 단계별 주요 내용 ····· 9
4. 장애인삶 패널조사 ····· 12
5. 3차년도 조사참여 현황 ····· 15

Ⅱ. 3차년도 무응답 특성 및 패널이탈 분석 ····· 21

1. 개요 ····· 23
2. 기준정의(용어) ····· 24
3. 표본유지율 및 조사완료 현황 ····· 26
   1) 표본유지율 ····· 26
   2) 연차별 조사완료 현황 ····· 27
4. 3차년도 무응답자 특성 분석 ····· 29
5. 응답패턴 분석 ····· 34
6. 소결 ····· 37

Ⅲ. 선행자료 분석 ····· **39**

1. 개요 ····· 41
2. 타 패널조사 분석 ····· 42
   1) 산재보험패널조사 ····· 42
      (1) 3차년 이후 종단가중치 조정 과정 ····· 42
      (2) 장애인삶 패널조사에 주는 시사점 ····· 43
   2) 한국청소년패널조사(KYPS) ····· 44
      (1) 3차년 이후 종단가중치 조정 과정 ····· 44
      (2) 장애인삶 패널조사에 주는 시사점 ····· 45
   3) 여성가족패널조사 ····· 46
      (1) 3차년 이후 종단가중치 조정 과정 ····· 46
      (2) 장애인삶 패널조사에 주는 시사점 ····· 49

Ⅳ. 장애인삶패널 가중치 조정에 대한 고찰(3차년도) ····· **55**

1. 기본방향 ····· 57
2. 3차년도 종단가중치 작성의 고려사항 ····· 58
3. 응답 여부와 주요 특성 변수들 간의 연관성 ····· 59
   1) 분석대상 ····· 59
   2) 패널 개인 및 가구특성 ····· 62
   3) 금융특성 ····· 64
   4) 영향요인 분석 ····· 65
4. 3차년도 가중치 처리 방안 비교·검토 ····· 67
   1) 산출하는 종단가중치의 종류 ····· 68
   2) 종단가중치(3차) 처리를 위한 검토 방안 ····· 70
   3) 검토 방안별 비교·검토 ····· 74
5. 주요항목 분석 ····· 81

Ⅴ. 결론 ····· **83**

참고문헌 ····· **87**

표목차

〈표 Ⅰ-2-1〉 장애인삶 패널조사의 응답 ····· 5
〈표 Ⅰ-2-2〉 패널조사 참여와 종단가중치 ····· 6
〈표 Ⅰ-3-1〉 연구 추진 절차 ····· 11
〈표 Ⅰ-4-1〉 조사영역 및 세부영역 ····· 14
〈표 Ⅰ-5-1〉 장애유형별 조사참여 현황 ····· 15
〈표 Ⅰ-5-2〉 장애정도별 조사 참여 현황 ····· 16
〈표 Ⅰ-5-3〉 성별 조사 참여 현황 ····· 17
〈표 Ⅰ-5-4〉 연령(패널) 조사참여 현황 ····· 17
〈표 Ⅰ-5-5〉 패널의 가구주 여부에 따른 조사참여 현황 ····· 18
〈표 Ⅰ-5-6〉 혼인상태별 조사참여 현황 ····· 18
〈표 Ⅰ-5-7〉 주택점유형태별 조사참여 현황 ····· 19
〈표 Ⅰ-5-8〉 가구원규모별 조사참여 현황 ····· 19
〈표 Ⅰ-5-9〉 가구소득별 조사참여 현황 ····· 20
〈표 Ⅱ-1-1〉 장애인삶 패널조사의 응답 ····· 24
〈표 Ⅱ-2-1〉 패널 분류 ····· 25
〈표 Ⅱ-3-1〉 조사완료 현황 ····· 26
〈표 Ⅱ-3-2〉 연차별 조사참여 여부에 따른 응답패턴 ····· 28
〈표 Ⅱ-3-3〉 연차별 조사참여 및 미참여 현황 ····· 28
〈표 Ⅱ-4-1〉 조사차수별 조사참여 및 조사미참여 현황 ····· 29
〈표 Ⅱ-4-2〉 3차 조사 기준 표본이탈(조사미참여) 사유별 현황 II-세부현황 ····· 31
〈표 Ⅱ-4-3〉 조사차수별 조사참여 및 조사미참여 현황 ····· 32
〈표 Ⅱ-5-1〉 2차-3차 조사참여 현황 ····· 34
〈표 Ⅱ-5-2〉 응답패턴 ····· 35
〈표 Ⅱ-5-3〉 응답패턴3 ····· 36
〈표 Ⅲ-2-1〉 산재보험패널조사 종단가중치 현황 ····· 43
〈표 Ⅲ-2-2〉 2차년도 응답자 중 3차년도 응답자 응답확률 추정결과-
여성가족패널조사 ····· 46

〈표 Ⅳ-2-1〉 가중치 부여 대상 ····· 58
〈표 Ⅳ-3-1〉 분석대상 ····· 61
〈표 Ⅳ-3-2〉 분석대상(2차-3차 조사 결과 비교) ····· 61
〈표 Ⅳ-3-3〉 패널의 개인 및 가구 특성 분석 결과 ····· 63
〈표 Ⅳ-3-4〉 패널의 금융특성 분석 ····· 64
〈표 Ⅳ-4-1〉 가중치 부여 대상 ····· 67
〈표 Ⅳ-4-2〉 응답패턴별 산출하는 종단가중치의 종류 ····· 69
〈표 Ⅳ-4-3〉 3차조사 결과에 따른 산출하는 종단가중치의 종류 및 부여 대상 ··· 69
〈표 Ⅳ-4-4〉 (방안1) 가장 최근가중치 승계 ····· 72
〈표 Ⅳ-4-5〉 (방안2) 2차년도 응답자의 3차년도 응답여부 기준으로 응답확률 추정 ····· 72
〈표 Ⅳ-4-6〉 종단가중치(3차) 산출 방안 검토 ····· 73
〈표 Ⅳ-4-7〉 방안2의 설명변수 ····· 74
〈표 Ⅳ-4-8〉 방안2의 응답확률 추정 모형(로지스틱 회귀) ····· 75
〈표 Ⅳ-4-9〉 방안3의 설명변수 ····· 76
〈표 Ⅳ-4-10〉 방안3의 응답확률 모형(로지스틱회귀모형) ····· 76
〈표 Ⅳ-4-11〉 검토방안별 종단가중치의 및 가중치 분포 ····· 78
〈표 Ⅳ-4-12〉 3차년도 종단가중치(wt3_1) 기술통계량_방안2 ····· 79
〈표 Ⅳ-4-13〉 3차년도 종단가중치(wt3_2) 기술통계량_방안2 ····· 79
〈표 Ⅳ-5-1〉 변수유형별 주요 문항 ····· 81
〈표 Ⅳ-5-2〉 주요문항별 상대표준오차 ····· 82

그림목차

[그림 Ⅱ-3-1] 주차별 조사완료 표본(2차(2019년), 3차(2020년) ····· 27
[그림 Ⅱ-4-1] 연도별 조사미참여(조사가능 中) 사유 비교 ····· 30
[그림 Ⅱ-4-2] 장애인삶 패널조사'조사가능' 및 '조사불가' 도표(3차조사) ····· 30
[그림 Ⅱ-4-3] 연차별 조사참여 및 미참여 ····· 33
[그림 Ⅱ-4-4] 연차별 조사미참여 인원: 조사가능 및 조사불가 ····· 33
[그림 Ⅱ-5-1] 응답패턴2(397명)-3차 조사 미참여 사유 ····· 36
[그림 Ⅳ-3-1] 개인 및 가구 특성의 CHAID 분석결과 ····· 63
[그림 Ⅳ-3-2] 금융특성의 CHAID 분석결과 ····· 65
[그림 Ⅳ-3-3] CHAID 분석결과 ····· 66
[그림 Ⅳ-4-1] 검토방안별 가중치 분포 ····· 78
[그림 Ⅳ-4-2] 3차년도 종단가중치 분포 wt3_2 (5,259명) ····· 80

한국장애인개발원
www.koddi.or.kr

연구요약

Ⅰ. 서론

1. 개요

○ 장애인삶 패널조사는 3차 조사 완료에 따라 이탈과 재진입한 표본이 발생함. 이에 시간의 흐름에 따른 유한모집단의 동태적 분석에 필요한 가중치와 당해 연도 조사 자료를 분석하는데 필요한 가중치 두 가지의 가중치 산출이 요구됨.

○ 이에 본 연구에서는 장애인삶 패널조사의 종단분석에 필요한 3차년도 종단가중치 작성을 위한 연구를 수행함. 3차년도에 재진입한 조사참여자의 특성과 1~3차 연속 참여자의 특성을 분석하고 종단가중치 작성시 이슈가 되는 사항을 파악하여 종단가중치 산출 방안을 도출함

○ 이를 통해 무응답자에 대한 보정가중치로써 시간의 흐름에 따른 조사 차수간의 비교 분석을 통한 유한모집단의 동태적 분석이 가능하도록 기반을 마련하고자함

2. 연구필요성 및 목적

○ 장애인삶 패널조사는 3차년도 조사에서 2차년도 조사에 참여하지 않았으나 3차년도 조사에 재진입한 새로운 응답패턴이 발생함. 이 경우는 이전년도 종단가중치가 존재하지 않는다는 문제를 가짐

○ 이에 3차년도 종단가중치를 도출함에 있어 1~3차 조사에 모두 참여한 연속참여자와 2차년도 탈락 후 3차년도에 재진입한 조사참여자를 고려한 종단가중치 산출 방안 마련이 필요함

○ 본 연구에서는 3차년도 종단가중치 산출을 위한 방안을 도출하는것을 목적으로 함. 이에 3차년도 재진입 표본에 대한 분석과 1~3차 년도 조사에서 나타나는 이탈 특성 분석을 실시하고자 함

○ 이러한 표본 이탈 특성 분석을 통해 장애인삶 패널조사의 표본이탈을 최소화하여 데이터 품질 관리와 표본유지율을 제고하는데 필요하 기초자료로 활용하고자함

3. 연구내용 및 방법

1) 연구내용

○ 3차년도(2020) 조사참여자 특성 분석
  - 3차 조사에 참여하지 않은 이탈 패널의 특성 및 무응답 특성 분석

○ 타 패널조사의 3차년도 이후 종단가중치 산출 과정 검토 및 분석
  - 장애인삶 패널조사의 3차년도 종단가중치 작성시 발생할 수 있는 주요이슈를 전반적으로 검토함. 조사참여자의 특성과 이에 수반되어 발생하는 종단가중치 작성시 이슈사항을 전반적으로 파악하고 정리하는 것을 목적함 표본추출, 1차 년도 가중치, 종단가중치 작성과정 분석

○ 3차년도 응답표본과 무응답표본 특성 분석, 최근가중치 승계, 무응답 보정을 위한 응답확률 추정 모형 검토
  - 조사참여에 영향을 미치는 주요 특성 변수 탐색. 무응답 표본의 가구 특성과 응답여부에 따른 패널 개인 및 가구, 금융변수와의 연관성을 파악하는 것을 중심으로 분석함. 이를 통해 3차년도 종단가중치 설계를 위한 기본 방향 설정
  - 최근가중치 승계 및 응답확률 추정 모형을 토대로 3차년도 가중치 작성 방안을 도출하고, 방안별 결과의 비교·검토를 통해 3차년도 가중치 작성 방안을 제시

2) 연구방법

○ 3차 조사 표본유지율 및 응답패턴 분석, 조사 참여 현황 분석, 장애인삶 패널조사에서의 이탈 패널 정의(용어)

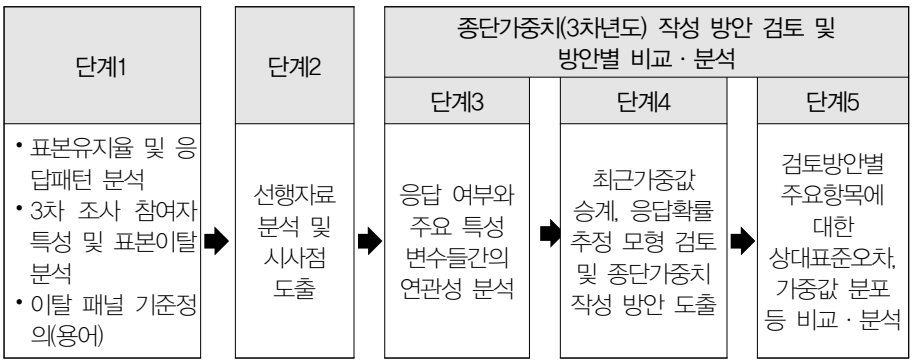
○ 타 패널조사 및 선행 자료의 3차년도 이후 종단가중치 산출 과정 검토, 전문과 자문회의 실시

○ 3차년도 조사참여자 특성 분석 및 응답여부와 따른 주요 특성변수의 연관성 분석

○ 3차 조사 참여 결과 기준 재진입 표본을 고려한 종단가중치 산정을 위해 최근가중값 승계, 응답확률 추정 모형 검토 및 산출 방향 설정

○ 검토방안별 주요항목에 대한 상대표준오차, 가중값 분포 등 비교·분석

[연구 수행 과정]



4. 3차년도 조사참여 현황

○ 3차년도 조사 참여 여부를 기준으로 1차 조사에 참여한 표본 중 3차 조사에서 이탈한 경우와 조사에 참여하여 응답한 경우로 나누어 분석함. 분석대상은 1차 조사 완료된 6,121명 패널을 기준으로 3차 조사 참여 완료한 표본은 5,259명, 조사미참여 표본은 862명임

○ 1차 조사에 응답 완료한 표본 중 3차 조사에서 이탈한 장애인 표본은 862명으로 14.1%의 표본이 이탈함

○ 종합해보면, 장애유형 중 내부·안면이 3차년도 조사 이탈 비중이 상대적으로 큼. 연령이 높은 집단에서 이탈율이 상대적으로 높음. 보다 구체적으로 살펴보면, 중증장애인의 이탈율이 경증에 비해 상대적으로 높음

Ⅱ. 3차년도 무응답 특성 및 표본이탈 특성 분석

1. 개요

○ 장애인삶 패널조사는 국내·외 많은 패널조사와 같이 초기(2차-3차) 이탈이 집중됨

○ 3차 조사 완료에 따라 3차년도 조사에서 2차 조사에 참여하지 않았으나 3차 조사에는 참여한 재진입 표본이 발생함

○ 본 연구에서 표본이탈 특성은 다음의 두 가지 목적을 가짐

- 첫 번째는 응답패턴과 이탈사유별 분석으로 주요 이탈 사유 분석을 통해 이탈 표본의 기본적인 특성 파악 및 기초 정보 확보
- 두 번째는 장애인삶 패널조사의 3차년도 조사의 응답패턴과 특성 변수 간의 연관성 분석을 실시함.

2. 기준정의(용어)

○ 본 연구에서 조사미참여자는 미참여 사유에 따라 구분함. 장기비수신, 조사거절(강력거절 포함), 주소불명, 기타는 '조사가능' 패널로 정의함. 사망, 장애등록취소, 시설입소, 병원입원 등은 '조사불가' 패널로 정의함

○ 조사불가 패널 중 사망, 장애등록취소는 향후 조사에 참여가 불가능하며 '소멸표본'으로 정의함

3. 표본유지율 및 조사완료 현황

○ 장애인삶 패널조사의 3차년도 조사완료 현황을 살펴보면, 1차 년도 구축된 장애인 6,121명 중 5,259명이 참여하였으며, 전체 응답률은 약 85.9%(=3차년도 조사 참여자/1차 년도 조사 참여자)임. 표본이탈율은 전체 14.1%로, 862명의 표본이탈 발생. 무응답이 전체의 약 9.3%이며, 소멸표본은 약 4.7%임

〈요약 표-1〉 조사완료 현황

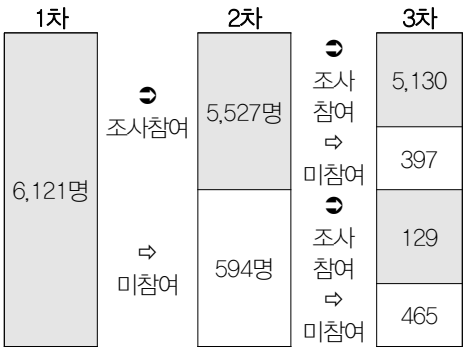
조사차수	조사완료 현황(명)			표본 유지율(%)	표본이탈	
	참여	미참여	계		(1차)응답 → (2차) 조사미참여(무응답)	조사불가 (2차)
1차(2018)	6,121	–	6,121	100	–	–
2차(2019)	5,527	594	6,121	90.3	471	123
3차(2020)	5,259주)	862	6,121	85.9	572	290

주) 원표본 유지율(%) = (5,259/6,121)×100

○ 장애인삶 패널조사의 경우 약 90%가 조사 초반에서 중반사이에 조사완료 되고 있음. 따라서 조사 초반에 패널 컨텍 및 참여 유도를 위한 방안 마련 필요

○ 3차 조사 참여 결과를 바탕으로 4가지 응답패턴으로 구분됨. '응답패턴3'은 3차 조사에서 처음 나타나는 응답패턴으로 3차 조사에서 재진입한 패널을 의미함

〈요약 표-2〉 연차별 조사참여 및 미참여 현황



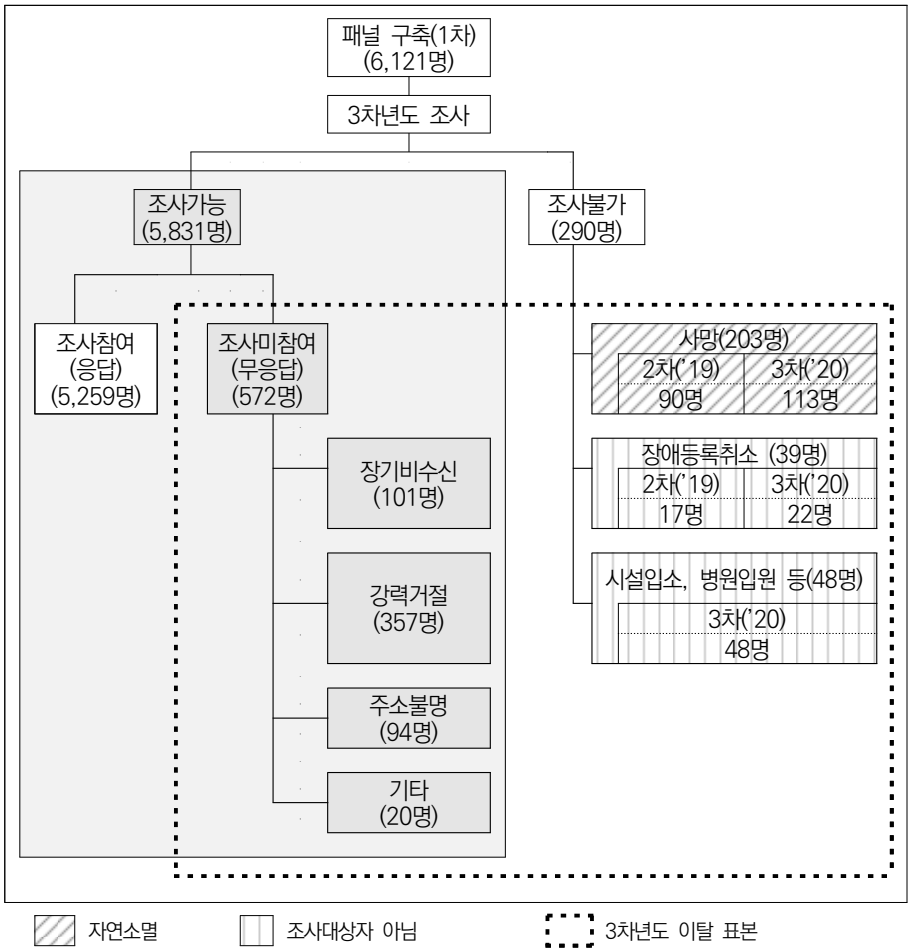
응답패턴	빈도(명)	비율(%)
111	5,130	83.81
110	397	6.49
101	129	2.11
100	465	7.60
	6,121	100.00

1: 조사참여
0: 조사미참여

4. 3차년도 무응답자 특성 분석

○ 본 연구에서는 이탈사유가 조사거부, 비수신 등으로 추후 차수 조사에 참여 가능성이 존재하는 경우를 무응답 표본으로 정의함. 사망, 이민 등조사가 불가능한 사유로 이탈한 경우를 소멸표본으로 정의하고 있음

○ 이탈 사유별 현황은 조사 강력거절이 41.4%로 가장 높게 나타났다. 다음으로 사망이 23.5%, 장기비수신이 11.7%, 주소불명 10.9% 순임



주) 김현지, 김태용, 이혜수. (2020).에서 발췌하여 3차년도에 맞게 재구성함. 출처: Hidiroglou et al. (1993). in Paul P. Biemer and Lars E. Lyberg. 2003. Introduction to Survey Quality, p.86.; 변종석, 이석훈, 정구현. (2013).에서 발췌하여 본 패널조사 내용으로 재구성함

[요약 그림-1] 장애인삶 패널조사 '조사가능' 및 '조사불가'도표(3차 조사)

○ ‘조사가능’ 패널의 미참여 사유 중 조사거절(강력거절 포함)이 가장 높은 비율을 차지함. 따라서 조사거절 패널에 대한 컨텍 및 참여 유도 방안에 대한 고민과 논의가 필요함

5. 응답패턴 분석

○ 장애인삶 패널조사 3차 조사 완료에 따른 응답패턴은 4가지로 구분됨
- 응답패턴1: 2차 조사에 참여한 5.527명 중 3차 조사에 참여한 패널로 1차~3차 조사에 모두 참여한 패널로 5,130명임
- 응답패턴2: 2차 조사 참여 후 3차 조사에 미참여한 경우로 397명임. 이 중 256명은 향후 조사 참여 가능성이 존재하는 ‘조사가능’ 패널이며, 141명은 사망, 장애등록취소 등 ‘조사불가’ 패널임
- 응답패턴3: 3차 조사에서 처음 발생한 패턴으로 2차조사에 미참여 후 3차조사에 참여한 패널로 129명임
- 응답패턴4: 1차 조사 이후 2,3차 조사에 미참여한 패널로 465명임

○ 이러한 결과를 종합해보면, 조사 미참여 사유가 조사거절(강력거절 포함), 장기비수신 등은 차년도 조사 참여를 유도할 수 있는 장치 마련이 필요할 것임

Ⅲ. 선행자료 분석

1. 개요

○ 본 절은 국내 · 외 패널조사의 가중치 산출 과정 분석을 통해 장애인삶 패널조사 3차년도 가중치 산출 방향 설정과 작성 방안 마련을 목적으로 함

2. 타 패널조사 분석

○ 타 패널조사를 살펴본 결과 장애인삶 패널조사 3차년도 가중값 산출을 위한 방향은 다음과 같음
- 3차년도 이후 종단가중치는 1차~3차에 모두 응답한 표본과 2차 조사 이탈 후 재진입한 표본을 고려한 두 가지 유형의 가중값 산출이 필요함

- 종단가중치 산출 시 3차년도 가중치 산출시 이전 년도와의 일관성 등의 측면을 고려하고 적절한 투입변수의 적용을 위한 사전 분석 실시 필요
- 응답확률 추정 모형의 설명변수는 응답 여부와의 연관성에 기반하여 응답에 영향을 줄 것으로 예상되는 설명변수를 적용할 필요가 있으며, 이를 위한 검토가 필요함

Ⅳ. 장애인삶패널 가중치 조정에 대한 고찰(3차년도)

1. 기본방향

○ 본 장에서는 패널의 표본 탈락을 고려한 장애인삶 패널조사의 3차년도 종단가중치 작성을 위한 다양한 방안을 검토함
- 장애인삶 패널조사의 경우 추출단위가 개인, 조사단위가 가구인 개인단위 패널조사 이므로 개인 패널의 표본 탈락 고려가 필요함

○ 1~3차 조사에 모두 참여한 연속참여 패널과 2차 조사에 이탈 후 3차 조사에 재진입한 패널의 상이한 특성을 고려하여 두 가지 유형의 가중값 산출이 필요함
- 이때 종단가중치는 1차년도부터 3차 조사에 모두 참여하고, 3차년도에 최종적으로 조사 완료된 5,259명에 대해 부여함
- 종단가중치는 무응답 조정한 가중치를 기본가중치로 정의하고, 레이킹 비조정하여 산출. 이때 표본설계 및 조사시점별 가중치 작성방법과의 일관성, 실제 공표와 향후 자료 활용 측면을 전반적으로 고려함

2. 3차년도 종단가중치 작성의 고려사항

○ 본 패널조사는 2차년도 조사에서 이탈 후 3차년도 조사에 재진입한 패널은 이전 차수의 가중치가 부재함

○ 따라서 3차년도 조사 결과에 따른 응답패턴 특성이 고려되어야 함
- 첫 번째는 1차~3차 모두 참여한 연속응답패널의 가중치 산출을 위한 적용 방안 검토 필요

- 두 번째는 2차 조사에 미참여 후 3차 조사에 재진입한 패널에 대한 가중치 처리 방안 검토 필요

〈요약 표-3〉 가중치 부여 대상

구분	계(명)	1차(2018)	2차(2019)	3차(2020)		
				응답여부	가중치 부여 대상	비고
응답패턴1	5,130	1	1	1	O	
응답패턴2	397	1	1	0	–	
응답패턴3	129	1	0	1	O	2차 가중치 없음
응답패턴4	465	1	0	0	–	
계(명)	6,121	6,121	5,527	5,259		

3. 응답 여부와 주요 특성 변수들 간이 연관성

○ 본 절에서는 3차년도 응답 표본과 무응답 표본 간이 차이를 특성 분석을 통해 파악하고 패널 이탈에 영향을 미치는 요인을 분석하고자 함

- 조사참여 여부를 종속변수로, 개인 및 가구 특성, 금융특성 관련 변수에 대해 카이제곱 검정(독립성), 연속형 변수의 경우 독립성 t검정 실시, 그리고 CHAID 분석 을 통해 응답여부에 영향을 미치는 주요 특성 변수를 파악함

○ 본 절에서의 분석대상은 1차-2차 조사에 참여한 5,527명 중 3차 조사 참여 여부에 따라 분류함. 이 때 이전 차수 가중치가 존재하지 않는 재진입 표본은 제외함

○ 장애등록취소, 사망 등으로 조사불가한 패널을 제외하고, 조사가능한 5,386명으로 정의

- 무응답자는 조사가능 하였으나 최종적으로는 조사 미참여한 패널로 비협조적 성향을 가지는 표본임

〈요약 표-4〉 분석대상

2차 조사 결과	3차조사 결과	인원(명)
조사 참여 (5,527)	참여(협조성향)	5,130
	미참여(비협조성향)	256
	미참여(사망, 장애등록취소, 시설입소 등(조사제외))	141
조사 미참여 (594)	참여(협조성향)	129
	미참여(비협조성향)	316
	미참여(사망, 장애등록취소, 시설입소 등(조사제외))	149
6,121	–	6,121

2차 조사 결과	3차 조사 결과	
	분석1 (wt3_1)	분석2 (wt3_2)
분석대상	분석대상	분석대상
분석대상	분석대상	분석대상
분석대상		
분석대상		분석대상
분석대상		분석대상
분석대상 (26명)		
5,998	5,386	5,831

⇧

응답확률 추정시 모형에 포함

○ 이 때 특성 변수는 설계변수, 공표단위, 1차 년도 가중치 조정 변수 등을 종합적으로 검토함

○ 분석을 통해 조사참여 여부와 패널에 대한 주요 특성 변수들간의 연관성 파악 및 최종 응답여부인 조사참여에 영향을 미치는 주요 특성 변수 검토

- 응답표본과 무응답표본에 대해 개인·가구 특성 변수로 패널의 성별, 연령대, 혼인 상태, 교육수준, 가구주 여부 등의 변수와 금융변수인 자산, 부채, 소득분포 비교
- 응답여부에 따른 특성변수와의 연관성 분석은 범주형 변는 카이제곱 검정(독립성), 연속형 변수의 경우 독립성 t검정 실시, 그리고 CHAID 을 통해 응답여부에 영향을 미치는 주요 특성 변수 파악

4. 3차년도 가중치 처리 방안 비교 · 검토

○ 장애인삶 패널조사 3차년도 가중치 부여 대상은 3차년도 조사에 응답 완료한 패널임

- 3차년도 응답 결과에 따라 1차~3차 조사에 모두 참여한 연속응답 패널과 3차년도 조사에 재진입한 패널이 가중치 부여 대상임

○ 본 절에서는 첫째, 재진입 패널을 어떻게 처리할 것인지, 둘째, 연속참여 패널의 가중치 산출을 위해 어떤 방안을 적용할 것인지에 대해 살펴보고자 함

○ 장애인삶 패널조사에서는 2가지 유형의 종단가중값을 산출하고자 하며, 하나는 1차~3차 조사에 모두 참여한 연속참여 패널에게 부여하는 가중치(wt3_1)이고 다른 하나는 3차년도 조사에 참여한 모든 패널에게 부여하는 가중치(wt3_2)임

〈요약 표-5〉 응답패턴별 산출하는 종단가중치의 종류

기준모집단: 패널구축시 모집단 N=251,277

가중치 유형 ①: wt3_1 (연속응답자)

응답패턴	1차(2018)	2차(2019)	3차(2020)
패턴1	1	1	1
패턴2	1	1	0
패턴3	1	0	1
패턴4	1	0	0

→ 종단

종단가중치 부여	인원(명)	5,527	5,130 (wt3_1)
기준모집단	계(명)	251,277	251,277

가중치 유형 ②: wt3_2 (해당차수 전체 응답자)

응답패턴	1차(2018)	2차(2019)	3차(2020)
패턴1	1	1	1
패턴2	1	1	0
패턴3	1	0	1
패턴4	1	0	0

↓ (기초 통계용)

종단가중치 부여	인원(명)	5,527	5,259 (wt3_2)
기준모집단	계(명)	251,277	251,277

○ 본 조사에서 가중치 산출을 위해 검토한 방안은 다음과 같음

- 방안1: 가장 최근가중치 승계
- 방안2: 2차년도 응답자의 3차년도 응답여부 기준 응답확률 추정
- 방안3: 방안2와 동일한 형태지만 3차 응답여부에 대한 응답확률을 추정하는 방안

〈요약 표-6〉 종단가중치(3차) 산출 방안 검토

산출 방안		종단가중치 유형 ①: wt3_1 (연속응답자: 5,130명)	종단가중치 유형 ②: wt3_2 (기초통계용: 5,259명)
최근 가중치 승계	방안 1	• 가장 최근 가중치를 원가중치로 무응답 보정 후 레이킹 비 조정 → 레이킹 비 조정(조정변수: 지역, 장애유형, 장애정도, 성별, 연령)	• 5,130명: 가장 최근 가중치(전년도 가중치)를 원가중치 • 129명(재진입): 가장 최근 가중값 승계 → 레이킹 비 조정(조정변수: 지역, 장애유형, 장애정도, 성별, 연령)
응답 확률 모형 + 최근 가중치 승계	방안 2	• 2차 응답자 기준으로 3차 응답여부에 대한 응답확률모형 적용(2차와 동일) → 레이킹 비 조정 – 응답확률모형: (설명변수) 지역, 장애유형, 장애정도, 성별, 연령, 주택점유형태, 가구월평균소득 – 레이킹 비 변수: 지역, 장애유형, 장애정도, 성별, 연령	• 5,130명: 2차 응답자 기준으로 3차 응답여부에 대한 응답확률모형 적용(2차와 동일) • 129명(재진입): 가장 최근 가중값 승계 → 레이킹 비 조정(조정변수: 지역, 장애유형, 장애정도, 성별, 연령)
응답 확률 모형	방안 3	• 3차 응답자 기준으로 3차 응답여부에 대한 응답확률모형 적용(2차와 동일) → 레이킹 비 조정 – 응답확률 추정 모형: (설명변수) 장애유형, 장애정도, 성별, 연령, 주택형태, 가구월평균생활비, 가구원수 – 1차 자료(모든 개체에 대해 값이 존재)	• 종속변수: 3차 응답여부 • 설명변수: 지역, 장애유형, 장애정도, 성별, 연령, 주택형태, 가구월평균생활비, 가구원 수 – 1차 자료(모든 개체에 대해 값이 존재)

○ 검토방안별 비교·검토를 위해 로지스틱 회귀분석을 적용하였고 설명변수는 앞 절에서 살펴본 응답자와 무응답자간 특성 분석 결과를 활용함
- 이때 조사 미참여 사유가 사망, 장애등록취소, 시설입소인 패널은 응답확률 추정 모형에서 제외함

○ 최종적으로 응답확률 추정 모형과 최근가중값 승계하는 방법을 혼합하는 형태인 방안2를 적용함

5. 주요항목 분석

○ 본 절에서는 1차 년도 가중치 산출시 검토 진행한 주요항목을 기준으로 상대표준오차를 산출하고, 검토방안별 비교·분석 실시

- 주요문항은 변수 유형별로 살펴보았으며, 본 패널조사 문항은 크게 연속형, 범주형 변수, 그리고 세부문항이 척도 응답을 받고 이를 평균하여 점수로 산출하는 문항임
- 연속형 변수는 월평균 가구소득, 노후생활에 필요한 1달 생활비, 척도 변수로 평균을 활용하는 장애수용, 단일 범주형 변수는 일상생활존중정도임. 이때 일상생활존중정도는 비율추정 변수임

○ 검토방안별 모집단 전체 및 관심 도메인 추정량의 상대표준오차는 유의한 차이를 보이지 않음. 일반적으로 허용되는 범위가 넘는 경우는 나타나지 않아 전반적으로 만족할 만한 수준인 것으로 판단됨

〈요약 표-7〉 주요문항별 상대표준오차

구분	wt3_1			wt3_2		
	방안1	방안2	방안3	방안1	방안2	방안3
월평균 가구소득	1.6644	1.6686	1.6664	1.6487	1.6532	1.6484
노후생활에 필요한 1달 평균 생활비	0.9752	0.9747	0.9761	0.9700	0.9694	0.9706
장애수용(평균)*	0.4107	0.4103	0.4101	0.4068	0.4065	0.4065
일상생활존중받는 정도(존중받고 있는 비율)**	1.7016	1.7025	1.7009	1.6878	1.689	1.6867

* 주 1) 12개 세부 항목의 전체 평균을 사용하여 산출(일부 문항 역코딩)
2) 4점 리커트 척도: 전혀 그렇지 않다. 거의 그렇지 않은 편이다. 그런 편이다. 매우 그렇다.

** 주 1) 일상생활존중받는 정도: 전혀 존중받지 못한다. 거의 존중받지 못하는 편이다. 존중받고 있는 편이다. 매우 존중받고 있다.
2) 존중받고 있는 비율은 '존중받고 있는 편이다'와 '매우 존중받고 있다'를 합친 비율이다.

한국장애인개발원
www.koddi.or.kr

Ⅰ. 서 론

1 개요

패널조사는 특정 시점 모집단을 대표하는 표본을 시간의 경과에 따라 추적하여 조사함으로서 모집단 전체의 평균적 특성 변화(순변동, net change)와 개인 수준의 특성 변화(총변동, gross changes)를 추정할 수 있다. 이는 집단의 역동성 연구에 유용하다는 장점이 있다 (박인호 외, 2020).[1] 이러한 종단연구는 태도나 의식의 변화와 같이 모집단의 변화양상에 대한 이해 또는 모집단 변화의 원인에 대한 인과적 설명을 목적으로 그 필요성이 증대되고 있다 (이경상 외, 2008)[2]. 그러나 다수의 시점을 조사함으로서 패널 추적 실패, 패널의 적격성 상실, 조사 미참여 현상으로 조사 차수가 경과하면서 패널 이탈이 증가하게 된다. 이러한 패널이탈은 패널조사의 다양한 어려움 중 하나이며, 패널이탈의 증가는 초기 구축한 표본의 대표성 훼손을 발생시켜 결과적으로는 추정의 편의를 야기할 가능성이 있다. 이에 탈락하는 개인 또는 가구의 특성을 파악하는 것은 무엇보다도 중요한 요소라고 할 수 있다(주재선 외, 2012).

이러한 어려움을 해결하는 하나의 통계적 방안은 조사시점 기준의 즉, 조사차수별 무응답을 보정하여 가중치에 반영하는 것이다. 유한모집단의 종단분석은 패널의 대표성과 추정 결과의 강건성 유지를 위해서는 종단가중치를 이용한 분석이 바람직할 것이다.[3] 패널조사는 이러한 통계적 접근과 함께 근본적으로 표본이탈 과정과 이탈 패턴에 대한 연구가 필수적이며(이경희와 민인식, 2015), 이를 바탕으로 이탈의 특성을 파악하고, 이탈 가능성을 파악하여 패널을 관리하는 것을 필요로 한다.

장애인삶 패널조사는 1차년도 조사와 패널구축이 동시에 진행되어, 2020년 3차 조사를 완료하였다. 이에 총 3개년의 조사 자료가 패널 데이터로 축적되었다. 패널구축이 동시에 진행되었던 1차년도 조사는 6,121명을 조사완료 하였다. 2차 조사에서는 다양한 사유로 인하 무응답 발생으로 총 5,527명 조사 완료하였다. 3차 년도 조사는 5,259명 조사 완료하였으며, 3차 조사부터 이전에는 발생하지 않았던 새로

1) 박인호 외. (2020). 이중틀 접근을 통한 사업체패널조사의 표본추가 및 가중치 산출. 조사연구. 21(3), 37-59.
2) 이경상 외. (2008). 한국 청소년 패널조사(KYPS) VI. 한국청소년정책연구원.
3) 박민규, 김사라. (2015). 한국노동패널조사 자료의 분석을 위한 패널 가중치 산출 및 사용방안 사례 연구. 조사연구. 16(1), 49-71.

운 특성이 나타난다. 2차년도 조사에 참여하지 않았으나 3차 년도 조사에 재진입(re-entrance)한 표본의 발생이다. 즉 패널조사의 특성상 이탈과 재진입을 반복하는 표본이 발생하게 된다. 이는 동일 대상을 반복적으로 조사하는 특성에 따른 것으로 이러한 특성의 패널조사 자료를 활용한 분석을 위해서는 두 개 종류의 가중치 산출이 필요하다. 첫 번째는 시간 경과에 따른 유한모집단의 동태적 분석에 필요한 가중치이며, 두 개 년도 이상의 변화 분석을 위해서는 이러한 조사차수별 특성을 반영한 가중치가 필요하다. 두 번째는 당해 연도 조사 자료를 이용한 통계 산출 및 분석을 위한 가중치이다.[4] 전자의 가중치를 종단가중치라고 부른다.

패널조사 자료의 분석을 위한 이러한 가중치는 기본적으로는 일반조사의 가중치 산출과 마찬가지로 추출확률 및 응답확률 그리고 캘리브레이션(calibration) 과정을 거쳐 산출된다. 패널조사를 통해 수집된 자료를 위한 가중치 작성 방안과 관련한 연구로는 Kalton과 Brick(1994), Duncan(1995)이 대표적 이다. Kalton과 Brick(1994)은 가중치 산출에 대한 이론적 연구에 초점을 맞추고 있다. Duncan(1995)은 용이한 가중치 산출을 위한 가중치 산출과정을 제안하였다.[5] 본 연구에서는 패널이탈 특성 분석 및 장애인삶 패널조사의 자료 분석을 위한 3차년도 가중치 산출을 위한 방안을 도출하고자 한다.

4) 박민규, 김사라. (2015). 한국노동패널조사 자료의 분석을 위한 패널 가중치 산출 및 사용방안 사례 연구. 조사연구. 16(1), 49-71.

5) 박민규, 김사라. (2015). 한국노동패널조사 자료의 분석을 위한 패널 가중치 산출 및 사용방안 사례 연구. 조사연구. 16(1), 49-71.

2 연구필요성 및 목적

종단분석은 모든 조사 차수에 응답한 응답자 자료를 기본으로 한다(김재광, 2017).[6] 패널조사가 실시된 1차 년도를 기준으로 가중치가 산출되는 해의 모두 응답한 자료를 최종응답자로 간주하여 그들에 대해 매년 새롭게 산출이 필요하다. 따라서 종단가중치 부여 대상은 패널 구축 시점부터 종단가중치가 부여되는 차수의 조사년도까지 조사에 연속적으로 참여한 대상자가 되는 것이다.

장애인삶 패널조사는 1차 년도 조사가 패널 구축과 동시에 진행되었기 때문에 2차년도 조사에서 최초의 표본 이탈이 발생하였다. 따라서 2차년도 조사는 웨이브 무응답, 다시 말해 1차 년도 패널조사 참여 후 2차년도 조사에 참여하지 않은 응답자로 직전년도 가중치가 모두 존재한다. 그러나 3차년도 조사에서는 이전 조사 즉 2차 조사에는 참여하지 않고, 3차 조사에 참여하여 패널로 재진입한 새로운 형태의 응답패턴이 나타났다(case C). 따라서 3차년도 이후 종단가중치의 산출은 1~2차에서 나타나지 않았던 응답패턴(이탈패턴)에 대한 고려가 필요하다. 3차년도 조사의 가중치 조정 대상은 3차 년도까지 지속적으로 조사에 참여한 표본(패널), 그리고 2차에서는 조사에 참여하지 않았으나 3차년도 조사에 다시 참여하여 조사 완료한 표본이다. 이는 〈표 I-2-2〉의 P01과 P03에 해당한다. 그러나 후자의 경우에는(2차 조사에는 참여하지 않았으나 3차 조사에 재진입한 표본) 이전 차수의 종단가중치는 존재하지 않는다는 문제가 있다.

〈표 Ⅰ-2-1〉 장애인삶 패널조사의 응답

구분	유형구분	응답여부			case
		1차조사(2018)	2차조사(2019)	3차조사(2020)	
1	완전응답	✓	✓	✓	A
2	소멸패턴	✓	✓	–	B
3		✓	–	–	B
4	비소멸 패턴(재진입)	✓	–	✓	C

6) 김재광. (2017). 표본조사론. 자유아카데미.

〈표 Ⅰ-2-2〉 패널조사 참여와 종단가중치

ID	1차	2차	3차	종단가중치 여부
P01	1	1	1	2~3차 종단가중치
P02	1	0	0	3차년도 종단가중치 없음
P03	1	1	0	2차년도 종단가중치 있음/3차년도 종단가중치 없음
P04	1	0	1	2차년도 종단가중치 없음/3차년도 종단가중치 필요

주) 1: 조사참여, 0: 조사미참여

이에 본 연구에서는 3차년도 재진입 표본(이들 표본은 2차년도 이탈로 인해 2차년도 조사의 가중치가 부여되지 않았다.)에 대한 가중치 부여 방안을 검토하여, 3차년도 가중치 작성 방안을 마련하고자 한다. 또한 향후 보다 질 좋은 패널 데이터 자료를 제공하기 위해 패널 이탈 분석을 실시하고자한다. 종단가중치는 표본 이탈로부터 야기되는 추정의 편향에 대응하는 (Uhirg, 2008: 이상호와 이혜정, 2008에서 재인용; 홍민기와 최효미, 2014; 하형석 외, 2018)[7] 방법이지만, 이와 더불어 조사과정에서 양적·질적 이탈의 최소화 노력이 병행할 필요가 있으며 (이상호, 2005), 표본 이탈 특성을 파악하여 응답과 무응답간의 차이 분석을 통한 체계적 데이터 품질 분석에 기반을 두어 표본 관리가 필요하기 때문이다. 1~3차 조사에서 나타나는 패널 이탈 패턴의 특성에 따른 분석을 토대로 이탈 가능성을 파악하고, 근본적으로는 표본 이탈을 최소화하기 위한 기초 자료를 확보함으로서 표본유지율 제고에 기여하고, 데이터 품질 관리를 위한 기반을 마련하고자 한다.

7) 하형석 외. (2018). 한국아동·청소년패널조사IX. 한국청소년정책연구원.

3 연구내용 및 방법

1) 연구내용

(1) 개요

본 연구에서는 장애인삶 패널조사의 패널이탈 특성을 분석하고 3차년도 종단가중치 작성을 위한 방안을 검토하여, 본 조사의 3차년도 이후 종단가중치 산출의 이론적 토대를 마련하고자 한다.

연구의 주요 내용은 크게 세 가지 이다. 첫째, 3차년도 조사참여 결과를 기준으로 장애인삶 패널조사의 무응답 특성과 패널이탈 특성을 분석하고, 둘째, 타 패널조사의 3차년도 이후 종단가중치 작성 과정을 검토하여 장애인삶 패널조사의 종단가중치를 산출시 고려가 필요한 사항을 도출한다. 셋째, 3차년도 응답을 기준으로 3차년도 응답 여부와 주요 특성 변수들 간의 연관성 분석을 실시하고, 최근가중치 승계, 응답확률 추정 모형 등을 검토함으로서 종단가중치 작성 방안을 마련한다. 그리고 최종적으로 종단가중치 작성을 위해 적용할 방안을 도출하고자 한다.

본 연구에서는 장애인삶 패널조사 종단분석을 위한 3차년도 종단가중치 작성 방안을 마련하고, 향후 보다 질 좋은 패널데이터 자료를 제공하기 위한 데이터 품질 관리의 기본이 되는 패널 이탈 분석을 실시하고자 한다.

(2) 연구의 주요 내용

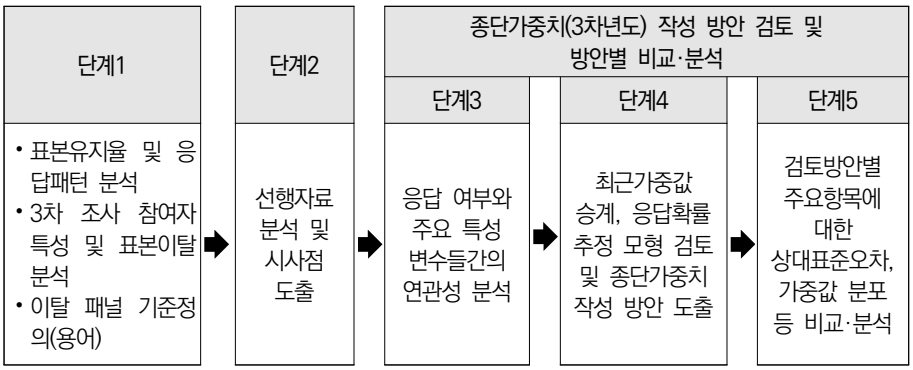
본 연구는 서론을 제외한 4개의 장으로 구성된다. 본 연구의 초점은 장애인삶 패널조사의 패널이탈 특성을 파악하여, 패널 관리의 기초자료 확보와 종단가중치 조정 방안을 마련하는 것이다. 이에 제2장은 응답(조사참여) 및 이탈 표본을 정의하고, 패널이탈 및 무응답 특성을 분석한다. 그리고 3차년도 조사참여 결과가 기준인 무응답 현황과 2차-3차 조사참여 결과를 기준으로 패널이탈 특성을 분석한다. 제3장에서는 장애인삶 패널조사의 종단가중치(3차년도) 산출을 위해 타 패널조사의 종단가중치 작성 과정을 중심으로 검토하여, 3차년도 가중치 작성시 발생할 수 있는 주요 이슈를 사전에 검토한다. 이는 벤치마킹할 수 있는 요인을 도출한다. 마지막

으로 4장에서는 3차년도 응답 여부와 주요 특성 변수들간의 연관성을 t-검정, 카이제곱검정, CHAID 분석을 통해 실시한다. 이를 통해 가구 특성과 금융특성 측면에서 응답여부와의 연관성을 파악한다. 그리고 3차년도 가중치 처리 방안을 도출하여 방안별 비교 · 분석한다. 이는 주요항목에 대한 상대표준오차, 가중치 분포 등을 검토한다. 이를 토대로 종단가중치 작성을 위한 방안을 제시하고, 5장에서 결론을 제시한다.

2) 연구 수행 방법

(1) 기본프레임

본 연구에서는 장애인삶 패널조사의 3차년도 종단가중치 산출을 위해 크게 3차년도 무응답 특성 및 패널이탈 특성 분석을 실시하고, 선행자료를 검토하는 문헌분석, 그리고 종단가중치 산출을 위한 다양한 방안을 검토하여 최종 방안을 제시한다. 표본이론과 실무적 접근에 있어서 발생하는 이슈는 학계 및 관련 전문가와의 논의를 토대로 결정하였다. 구체적인 연구 수행 방법은 ① 이탈 패널의 기준 정의(용어), 표본유지율 및 응답패턴 분석, 3차 조사 참여자 특성 및 표본이탈 분석 ⇒ ② 타 패널조사의 종단가중치(3차년도 이후) 산출 과정 검토 ⇒ ③ 3차년도 응답여부와 주요 특성 변수들간의 연관성 분석 ⇒ ④ 종단가중치(3차년도) 작성을 위한 최근가중값 승계, 응답확률 추정 모형 검토 ⑤ 종단가중치 작성 방안별 비교 · 분석의 5단계로 구성된다.



(2) 단계별 주요 내용

○ (단계1) 표본유지율 및 응답패턴 분석

1단계에서는 1~3차 패널조사의 조사완료 결과를 바탕으로 이탈 패널의 기준(용어)을 정의한다. 3차 조사 참여 결과 분석을 통해 표본유지율과 3차 조사 참여자 특성을 분석하고, 표본유지율과 표본이탈 분석을 실시한다.

○ (단계2) 선행자료 분석 및 시사점 도출

선행자료 분석은 국·내외 패널조사를 검토하여 종단가중치 산출의 작성과정 등을 검토하여, 장애인삶 패널조사의 종단가중치 산출을 위한 시사점을 도출한다. 본 패널조사의 종단가중치 산출시 참고할 수 있는 방법 등 가중치 산출시 발생할 수 있는 이론적·실무적 이슈를 사전에 검토하고자 한다.

○ (단계3) 패널이탈 특성 및 응답여부와 주요 특성 변수들간의 연관성 분석

장애인삶 패널조사는 2차년도 조사에서부터 단위무응답 즉, 표본이탈이 발생하며, 3차년도 조사에서는 재진입 패널이 나타난다. 재진입 패널이란 2차 조사 미참여 후 3차조사 참여한 표본(패널)을 의미한다. 3단계에서는 응답패턴 분석을 통해 패널이탈 특성을 살펴본다. 그리고 3차년도 무응답 패널과 응답 패널간의 차이를 파악하고자 특성 변수간 분석을 실시한다. 이를 토대로 패널 이탈에 영향을 미치는 요인을 파악하여, 조사참여에 영향을 미치는 특성 변수를 탐색한다. 조사 결과를 바탕으로 응답 패널과 무응답 패널간 이탈 특성 분석을 실시하여 패널 이탈 관리 및 데이터 품질 관리를 위한 기반을 마련 한다. 분석은 패널 및 가구 특성 변수와 금융특성 변수를 대상으로 실시하며, 조사참여 여부와 패널의 주요 특성 변수(개인 및 금융)간의 연관성 분석과 조사참여 여부에 영향을 미치는 변수를 파악한다. 이를 통해 종단가중치 설계에 필요한 방향과 무응답 조정 모형의 기초정보를 마련한다.

○ (단계4) 종단가중치(3차년도) 작성 방안 검토

패널조사의 종단가중치는 이전 차수 가중치를 기반으로 한다. 장애인삶 패널조사의 3차년도 조사 참여 결과를 기준으로 2차 조사에 참여한 표본은 이전 차수 가중치가 존재한다. 하지만 2차 조사에 미참여한 표본은 이전 차수인 2차 조

사에 대한 가중치가 없다. 따라서 패널의 응답패턴을 고려하여 종단가중치(3차년도) 작성 방안을 검토한다. 이는 가중치가 존재하지 않는 재진입(re-entrance) 표본에 대한 가중치 처리와 이전차수 가중치가 존재하는 표본에 대한 가중치는 어떻게 처리할 것인가에 대한 문제이다. 이를 위해 최근가중값 승계, 응답확률 추정 모형을 검토하고 장애인삶 패널조사를에 적합한 종단가중치 작성 방안을 마련한다.

○ (단계5) 검토방안별 주요항목에 대한 상대표준오차, 가중값 분포 등 비교 · 분석

선행자료 분석과 패널이탈 특성, 3차년도 조사의 응답여부와 주요 특성 변수간의 연관성 분석을 토대로 최종가중치 승계 방안과 응답확률 추정 모형을 검토하여 종단가중치(3차년도) 작성 방안을 마련한다. 검토방안별 가중치는 주요항목에 대한 상대표준오차와 가중값 분포 등을 통해 비교 · 분석한다. 분석결과를 기반으로 전문가 자문회의를 거쳐 이론적 · 실무적 검토 후 최종가중치 산출 방안을 제시한다.

〈표 Ⅰ-3-1〉 연구 추진 절차

연구계획수립 및 자문위원 구성	▸세부 연구계획 수립 ▸자문위원 구성
⬇	
선행 자료 검토	▸종단가중치관련 문헌검토 및 타 조사의 종단가중치 부여 방법(3차이후) 분석 – 가중치 작성 방법 및 작성시 주요 고려 사항 점검(방법론에 대한 이론 및 참고할 시사점 도출 등)
⬇	
패널이탈 분석	▸1~3차 패널이탈 패턴 분석 ▸응답패턴 및 용어정의
⬇⬆	
자문회의	▸종단가중치 조정을 위한 방안 수립
⬇⬆	
종단가중치 설계	▸종단가중치(3차년도) 작성 방안 마련 –최근가중값 승계, 응답확률 모형 검토 ▸주요문항에 대한 상대표준오차 산출 및 검토
⬇⬆	
자문회의	▸종단가중치 작성 결과 검토 및 의견 수렴 ▸패널이탈 및 무응답 특성 분석 결과 논의 ▸통계작성변경승인관련 사항 의견 수렴
⬇	
결과정리 및 보고서 작성	▸결과정리 및 보고서 작성

4 장애인삶 패널조사[8)]

1) 법적근거 및 통계작성승인번호

◦ 장애인삶 패널조사는 장애인복지법 시행령 제19조 제2항
◦ 통계법 제18조 및 같은 법 시행령 제24조: 승인번호(제438001호)[9)]

2) 조사연혁

◦ 장애인삶 패널조사 통계작성승인(작성승인일: 2018.9.6.)
◦ 1차 조사 및 패널구축 완료 (6,121명): 2018년
◦ 2차 조사 완료 (5,527명): 2019년
◦ 3차 조사 완료 (5,259명): 2020년

3) 조사주기 및 조사기간

◦ 조사주기: 1년, 조사기간: 매년 10월~12월

4) 조사목적

장애인삶 패널조사는 장애 발생에 따른 장애수용 과정과 사회관계 속에서 겪는 변화에 대해서 개인·가족·사회적 요인의 동태를 파악 하여, 향후 관련 정책의 수립·지원에 필요한 기초자료를 확보하는데 그 목적이 있다.[10)] 이에 장애등록이후 삶의 변화를 탐색할 수 있는 과학적이고, 체계적인 종단자료를 구축하고자 한다.

---

8) 본 절은「김현지, 김태용. (2019). 장애인삶 패널조사 가중치 산정 방안 연구. 한국장애인개발원」에서 발췌하였다.(p.4–6)
9) 통계청 고시 제2018–373호. 승인일: 2018.9.6.
10) 장애인삶 패널조사는 장애등록 이후 삶의 변화(장애수용, 사회복귀, 가족구조 등)를 파악하고, 삶과 관련된 일상생활, 소득수준, 건강실태, 복지욕구, 사회참여 등의 탐색을 통해 향후 관련 정책의 수립 · 지원에 필요한 자료를 마련하고자 한다. 이에 개인·가족·사회적 관점에서 장애 발생에 따른 장애수용과정과 사회관계의 변화 등을 탐색하고, 장애수용 및 변화, 건강·의료, 일상생활 등의 이해를 통해 관련 정책의 수립·지원 과정에서 활용 할 수 있는 기초자료를 확보하고자 한다.

5) 목표모집단

장애인삶 패널조사의 모집단은 장애인복지법 제2조에서 정의하는 장애인으로 같은 법 제32조에 의거하여 2015년 1월1일부터 2017년 12월 31일 사이에 보건복지부에 장애등록을 마친 등록장애인(단, 장애인거주시설 거주 장애인 제외)과 그들의 가구원이다.

6) 조사대상 및 조사방법

조사대상은 장애인복지법 제2조에서 정의하는 장애인으로 같은 법 제 32조에 의거 2015년 1월 1일부터 2017년 12월 31일 사이에 보건복지부에 장애등록 한 등록장애인(장애인거주시설 거주 장애인 제외)과 그들의 가구원이다.

조사대상은 표본설계[11]에 따라 표본배분 및 표본추출(장애유형, 장애정도, 성별 등) 하였다. 2018년(1차년도) 구축 완료된 패널 6,121명과 그들의 가구원을 대상으로 조사를 실시한다. 조사방법은 전문면접원에 의한 대면면접조사로 조사도구는 TAPI[12]를 활용한다.

7) 조사내용

장애인삶 패널조사는 장애인과 그들의 가구원이 장애수용 과정과 사회관계 속에서 겪는 변화를 '장애수용 및 가계변화, 건강·의료, 자립, 사회참여'라는 관점에서 접근하고자 하였으며, 영역별 세부 포함 내용[13]은 다음과 같다. 패널조사의 특성을 고려, 추적 조사가 실시되도록 마지막에 향후 이사의향 등과 같은 패널 관리 정보를 추가로 조사한다.

---

11) 김현지, 김태용. (2018). 장애인패널조사 표본설계. 한국장애인개발원.」을 참고한다.
12) TAPI : Tablet Assisted Personal Interviewing
13) 조사대상의 범위에 대한 자문회의(2017.11.28) 결과를 바탕으로 복건복지부와 협의를 통해 주요 조사 영역을 결정하였다. 이에 관한 자세한 내용은 장애인패널 예비조사(2018)를 참고한다.

〈표 Ⅰ-4-1〉 조사영역 및 세부영역

주요 조사 영역	세부 포함 내용
장애수용 및 변화	장애수용, 심리상태, 사회관계 경제상황, 주거, 일반현황 등
건강·의료	건강, 의료서비스, 안전
자립	일상생활, 여가활동, 자립생활, 교육, 고용(취업)
사회참여	장애관련서비스, 인권, 결혼생활 및 성, 노후, 삶의 만족도

주) 조사영역은 패널용(장애인용) 조사표 기준임

조사표는 패널용, 가구주용, 가구원용[14]으로 총 3종 이며, 각각의 조사내용은 다음과 같다.

〈표 1-4-2〉 조사표

장애인용 (패널)	가구주용	가구원용
표본으로 선정된 장애인(패널)에 관한 문항	표본으로 선정된 장애인의 가계변화에 관한 문항	표본으로 선정된 ① 패널에 관한 가구원의 경험·생각(인식)관련 문항 ② 가구원 본인과 관련한 문항
조사표 1	조사표 2	조사표 3

- 패 널 용 : 장애수용, 심리상태, 사회관계 및 건강, 의료서비스, 안전, 일상 및 여가생활, 자립생활, 교육, 고용(취업), 장애관련 서비스, 인권, 결혼생활 및 성, 노후, 삶의 만족도 등에 대하여 조사
- 가구주용 : 가구원 현황, 경제상황, 주거에 대하여 조사
- 가구원용 : 장애수용, 심리상태, 사회관계 및 건강, 의료서비스, 안전, 일상 및 여가생활, 자립생활, 교육, 고용(취업), 인권, 노후, 삶의 만족도 등에 대하여 조사

8) 통계 자료 산출 범위

본 패널조사 데이터를 활용한 자료 산출은 전국, 장애유형별, 장애정도별, 성별 주요 항목에 대한 통계 생산을 목표로 한다.

14) 가구원 설문의 경우 조사 참여 의향을 묻고, 참여 의향이 있는 경우만 응답 받는다. 가구원 조사표는 가구원 수만큼 조사한다.

5 3차년도 조사참여 현황

본 절에서는 3차년도 조사에 참여하지 않아 조사완료 되지 않은 패널을 조사참여 여부로 분류하여 분석한다.[15] 분석 대상은 1차 조사 완료된 6,121명을 기준으로 한다. 3차 조사 참여 여부를 기준으로 조사참여 표본은 5,259명, 조사미참여 표본은 862명으로, 3차 조사는 원표본 대비(1차 조사 완료) 85.9%의 표본유지율을 보였다. 3차 년도 이탈한 표본은 14.1%이다.

〈표Ⅰ-5-1〉은 장애유형에 따른 조사 참여 현황이다. 조사 미참여 비율은 안면장애가 26.67%로 타 장애 유형에 비해 높은 것으로 나타났다.

〈표 Ⅰ-5-1〉 장애유형별 조사참여 현황

구분	조사참여 여부		총계	행(열)%		총계(%)
	참여	미참여		참여	미참여	
지체	856	118	974	87.89 (16.28)	12.11 (13.69)	100 (15.91)
시각	654	103	757	86.39 (12.44)	13.61 (11.95)	100 (12.37)
청각	689	72	761	90.54 (13.10)	9.46 (8.35)	100 (12.43)
언어	333	47	380	87.63 (6.33)	12.37 (5.45)	100 (6.21)
지적	352	49	401	87.78 (6.69)	12.22 (5.68)	100 (6.55)
뇌병변	824	156	980	84.08 (15.67)	15.92 (18.10)	100 (16.01)
자폐성	146	20	166	87.95 (2.78)	12.05 (2.32)	100 (2.71)
정신	277	51	328	83.45 (5.27)	15.55 (5.92)	100 (5.36)
신장	484	96	580	83.45 (9.20)	16.55 (10.61)	100 (9.48)
심장	107	13	120	89.17 (2.03)	10.83 (1.51)	100 (1.96)

15) 본 절에서의 3차년도 조사참여 현황은 당해차수 응답현황으로 가중치를 적용하지 않은 수치로, 구축 패널 6,121명을 기준으로 조사참여 및 미참여 비율을 의미한다.

구분	조사참여 여부		총계	행(열)%		총계(%)
	참여	미참여		참여	미참여	
호흡기	121	42	163	74.23 (2.30)	25.77 (4.87)	100 (2.66)
간	167	39	206	81.07 (3.18)	18.93 (4.52)	100 (3.37)
안면	22	8	30	73.33 (0.42)	26.67 (0.93)	100 (0.49)
장루 · 요루	119	29	148	80.41 (2.26)	19.59 (3.36)	100 (2.42)
뇌전증	108	19	127	85.04 (2.05)	14.96 (2.20)	100 (2.07)
총계	5,259	862	6,121	85.92 (100)	14.08 (100)	100 (100)

다음은 장애정도별 조사 참여 현황이다. 3차년도 조사에 참여한 비중을 살펴보면, 중증이 전체의 51.97%, 경증이 48.03%이다. 중증 장애인의 조사미참여 비율은 경증에 비해 상대적으로 높은 것으로 나타났다.

〈표 Ⅰ-5-2〉 장애정도별 조사 참여 현황

구분	조사참여 여부		총계	행(열)%		총계(%)
	참여	미참여		참여	미참여	
중증	2,733	509	3,242	84.30 (51.97)	15.70 (59.05)	100 (52.97)
경증	2,526	353	2,879	87.74 (48.03)	12.26 (40.95)	100 (47.03)
총계	5,259	862	6,121	85.92 (100)	14.08 (100)	100 (100)

성별에 따른 조사참여 여부를 살펴보면, 3차 조사 년도 참여자 중 남자의 비중이 55.47%, 여자가 44.53%이다. 성별에 따른 조사 미참여 비율은 남자가 13.93%, 여자가 14.28%로 큰 차이를 보이지 않는다.

〈표 Ⅰ-5-3〉 성별 조사 참여 현황

구분	조사참여 여부		총계	행(열)%		총계(%)
	참여	미참여		참여	미참여	
남자	2,917	472	3,389	86.07 (55.47)	13.93 (54.76)	100 (55.37)
여자	2,342	390	2,732	85.72 (44.53)	14.28 (45.24)	100 (44.63)
총계	5,259	862	6,121	85.92 (100)	14.08 (100)	100 (100)

패널 구축 당시를 기준으로 연령대별 조사참여는 〈표 Ⅰ-5-4〉이다. 3차 조사에 참여하지 않은 비율은 만18세 이하가 13.23%, 만19세-49세 이하가 28.07%, 만50세 이상이 58.70%이다. 만19세-49세 이하에서 3차년도 조사 미참여 비율이 상대적으로 높은 것으로 나타났다.

〈표 Ⅰ-5-4〉 연령(패널) 조사참여 현황

구분	조사참여 여부		총계	행(열)%		총계(%)
	참여	미참여		참여	미참여	
만18세 이하	933	114	1,047	89.11 (17.74)	10.89 (13.23)	100 (17.11)
만19세-49세 이하	1,251	242	1,493	83.79 (23.79)	16.21 (28.07)	100 (24.39)
만50세 이상	3,075	506	3,581	85.87 (58.47)	14.13 (58.70)	100 (58.50)
총계	5,259	862	6,121	85.92 (100)	14.08 (100)	100 (100)

주) 연령은 2018년(1차 조사) 기준임

패널의 가구주 여부에 따른 조사참여 현황을 보면, 패널이 가구주인 경우는 48.91%, 가구주가 아닌 경우 51.09%이다. 패널이 가구주인 경우와 가구주가 아닌 경우에 따른 조사 참여 현황은 큰 차이를 보이지 않았다.

〈표 Ⅰ-5-5〉 패널의 가구주 여부에 따른 조사참여 현황

구분	조사참여 여부		총계	행(열)%		총계(%)
	참여	미참여		참여	미참여	
가구주	2,572	434	3,006	85.56 (48.91)	14.44 (50.35)	100 49.11
가구주아님	2,687	428	3,115	86.26 (51.09)	13.74 (49.65)	100 50.89
총계	5,259	862	6,121	85.92 (100)	14.08 (100)	100 (100)

다음은 패널의 혼인상태에 따른 조사 참여 현황이다. 기혼 45.56%, 미혼 33.14%, 이혼·별거·사별이 21.30%로 나타났다. 이혼 · 별거 · 사별인 경우가 기혼이나 미혼에 비해 이탈 비율이 상대적으로 높은 것으로 나타났다.

〈표 Ⅰ-5-6〉 혼인상태별 조사참여 현황

구분	조사참여 여부		총계	행(열)%		총계(%)
	참여	미참여		참여	미참여	
기혼(사실혼, 동거 포함)	2,396	391	2,787	85.97 (45.56)	14.03 (45.36)	100 (45.53)
미혼(만19세 미만 포함)	1,743	272	2,015	86.50 (33.14)	13.50 (31.55)	100 (32.92)
이혼 · 별거 · 사별	1,120	199	1,319	84.91 (21.30)	15.09 (23.09)	100 (21.55)
총계	5,259	862	6,121	85.92 (100)	14.08 (100)	100 (100)

주택점유형태에 따른 3차년도 조사 참여 현황을 보면, 3차년도 조사에 참여한 패널(표본) 중 자가라고 응답한 비율이 52.42%, 전세 등이 47.58%이다. 전세 등의 경우가 3차년도 조사 미참여 비율(15.01%)이 상대적으로 높은 것으로 나타났다.

〈표 Ⅰ-5-7〉 주택점유형태별 조사참여 현황

구분	조사참여 여부		총계	행(열)%		총계(%)
	참여	미참여		참여	미참여	
자가(자기집)	2,757	420	3,177	86.78 (52.42)	13.22 (48.72)	100 (51.90)
전세 등	2,502	442	2,944	84.99 (47.58)	15.01 (51.28)	100 (48.10)
총계	5,259	862	6,121	85.92 (100)	14.08 (100)	100 (100)

다음으로 가구원규모에 따른 3차 조사 참여 현황이다. 1인가구 17.91%, 2인 가구는 32.15%, 3인가구의 경우 22.68%, 4인가구 이상이 27.25%로 나타났다. 가구원규모별 조사 미참여 비율은 1인가구가 16.49%로 나타났다.

〈표 Ⅰ-5-8〉 가구원규모별 조사참여 현황

구분	조사참여 여부		총계	행(열)%		총계(%)
	참여	미참여		참여	미참여	
1인가구	942	186	1,128	83.51 (17.91)	16.49 (21.58)	100 (18.43)
2인가구	1,691	277	1,968	85.92 (32.15)	14.08 (32.13)	100 (32.15)
3인가구	1,193	187	1,380	86.45 (22.68)	13.55 (21.69)	100 (22.55)
4인가구 이상	1,433	212	1,645	87.11 (27.25)	12.89 (24.59)	100 (26.87)
총계	5,259	862	6,121	85.92 (100)	14.08 (100)	100 (100)

다음은 가구소득에 따른 3차 조사 참여 현황이다. 3차 조사 참여자의 가구소득별 분포를 살펴보면, 100만원 미만인 경우가 21.45%, 100만원 이상-200만원 미만인 경우가 21.64%, 200만원이상-300만원 미만인 경우가 18.31%, 300만원이상-400만원 미만인 경우가 15.44%, 400만원 이상인 경우가 23.16%로 나타났다. 100만원 미만인 경우가 가구소득별 조사 미참여 비율이(16.13%) 상대적으로 높은 것으로 나타났다.

〈표 Ⅰ-5-9〉 가구소득별 조사참여 현황

구분	조사참여 여부		총계	행(열)%		총계(%)
	참여	미참여		참여	미참여	
100만 미만	1,128	217	1,345	83.87 (21.45)	16.13 (25.17)	100 (21.97)
100만 이상-200미만	1,138	201	1,339	84.99 (21.64)	15.01 (23.32)	100 (21.88)
200만 이상-300미만	963	148	1,111	86.68 (18.31)	13.32 (17.17)	100 (18.15)
300만 이상-400미만	812	105	917	88.55 (15.44)	11.45 (12.18)	100 (14.98)
400만 이상	1,218	191	1,409	86.44 (23.16)	13.56 (22.16)	100 (23.02)
총계	5,259	862	6,121	85.92 (100)	14.08 (100)	100 (100)

종합해보면, 안면 장애인의 3차년도 미참여 비율이 가장 높은 것으로 나타났으며, 중증장애인, 연령대가 높은 집단, 소득이 낮은 집단, 주택점유형태가 전세 등인 경우, 1인 가구인 경우가 3차년도 조사미참여 비율이 상대적으로 높게 나타났다. 패널의 성별, 패널의 가구주 여부에 따른 조사참여 및 미참여 비율은 큰 차이를 보이지 않았다.

한국장애인개발원
www.koddi.or.kr

Ⅱ. 3차년도 무응답 특성 및 패널이탈 분석

1 개요

장애인삶 패널조사의 조사대상은 1차 년도 구축한 6,121명이다. 〈표 Ⅱ-1-1〉과 같이 2차년도 조사부터 패널조사의 특수한 형태 무응답인 단위무응답 즉, 표본이탈이 발생하며. 3차년도 조사에서는 재진입 패널이 나타난다. 재진입 패널이란 2차 조사 미참여 후 3차조사 참여한 표본(패널)이다. 일반적으로 패널조사는 조사차수가 증가하면서 패널 이탈률이 높아질 가능성이 크고, 이로인해 추정결과의 편의도 야기할 수 있다. 국내 · 외 많은 패널조사를 살펴보면, 특히 2차-3차조사에서 패널이탈이 집중적으로 나타나는 것으로 알려져 있다. 장애인삶 패널조사 역시 동일한 현상이 나타나고 있다. 이러한 문제에 대한 근본적 해결책은 패널(표본)이탈을 줄이는 것이며, 이를 위해서는 무응답 특성 및 패널이탈 과정과 패턴에 대한 정확한 이해가 선행적으로 필요하다.[16)]

장애인삶 패널조사는 추출단위가 개인이며, 조사단위가 가구인 패널조사이다. 이에 장애인 개인을 조사대상으로, 함께 거주하는 가구원을 포함하여 조사하고 있다. 따라서 이탈하는 패널 개인의 특성, 즉 이탈 장애인의 특성을 파악하는 것이 중요하다고 할 수 있다. 이에 패널 이탈 사유 분석을 토대로 이를 관리하는 것은 중요한 과제이다.

본 연구에서는 패널의 이탈 특성 파악을 위해 두 가지 측면에서 접근하고 한다. 첫 번째는 응답패턴 및 이탈사유별 분석이다. 응답거절, 사망 등 다양한 사유로 패널이탈이 발생하여 표본유지율에 영향을 주고 있다. 이에 본 장에서는 응답패턴과 이탈 사유별 분석을 토대로 주요 이탈 사유를 분석하여 조사관리에 필요한 기초정보를 확보하고자 한다. 두 번째는 어떤 요인이 패널 이탈에 영향을 미치는지에 대한 것이다. 이에 대해서는 Ⅳ장에서 살펴본다. 이를 위해 1차-3차 조사참여 결과를 토대로 응답 및 이탈 표본에 대해 정의한다.

16) 이경희, 민인식. (2016). 패널조사 응답지속성에 관한 연구: 한국노동패널조사를 중심으로.

〈표 Ⅱ-1-1〉 장애인삶 패널조사의 응답

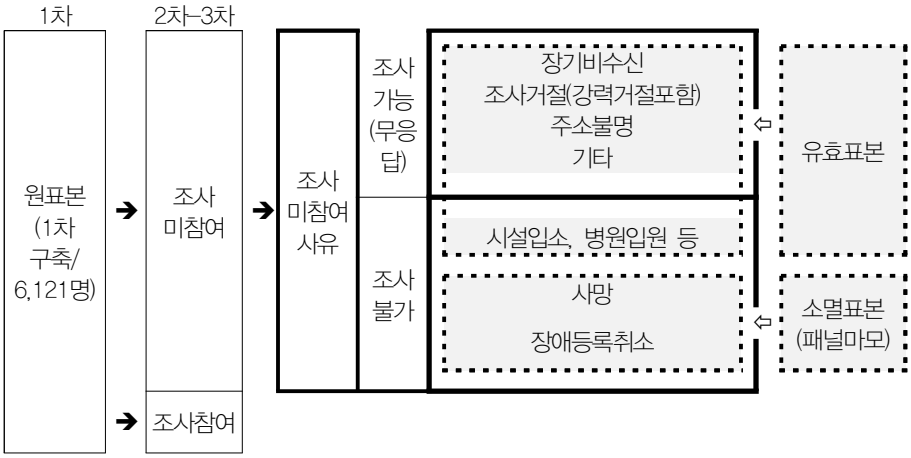
번호	유형구분	응답여부		
		1차 조사 (2018)	2차 조사 (2019)	3차 조사 (2020)
1	완전응답	✓	✓	✓
2	소멸패턴	✓	✓	–
3		✓	–	–
4	비소멸 패턴(재진입)	✓	–	✓
응답 및 응답률		6,121	5,527	5,259

출처: 「김규성, 박인호. (2010). 패널조사 웨이브 무응답의 대체방법 비교」에서 발췌하여, 장애인삶 패널조사에 맞게 재구성함.

2 기준정의(용어)

본 연구에서 응답 및 이탈 표본 정의는 다음과 같다. 조사미참여자는 조사 미참여 사유에 따라 조사가능하였으나 미참여한 패널과 조사불가 사유로 미참여한 패널로 구분할 수 있다. 조사참여 여부에 따른 패널의 특성은 조사미참여 사유에 따라 다를 수 있으며, 조사에 미참여한 패널 중에서 장기비수신, 주소불명, 기타는 "조사가능" 패널로 정의하였다. 시설입소, 병원입원 등과 사망, 장애등록취소는 당해 조사불가 패널로 정의되며, 특히 이들 중 향후 차수의 조사에는 참여하지 못하는 소멸표본(사망, 장애등록취소)으로 정의하였다.

〈표 Ⅱ-2-1〉 패널 분류



◦ 패널 마모: 조사미참여 사유가'사망', '장애등록취소'인 표본

◦ 원표본: 1차 년도 구축된 패널 6,121명

(* 장애인삶 패널조사는 패널구축 이후 표본 추가가 이뤄지지 않았으며, 1차 년도 구축된 패널을 매년 추적 조사함)

◦ 소멸표본(패널마모): 조사미참여 사유가 '사망', '장애등록취소' 패널

◦ (당해년도) 유효표본: 이후 조사 참여가 불가능한 자연소멸(사망)과 장애등록취소를 제외한 패널 ➜ 패널 유지율 제고를 위한 관리 대상 패널

◦ 조사불가: 당해연도 조사가 불가능한 표본으로 조사미참여 사유가 '사망', '장애등록취소', '시설입소 · 병원입원 등'임. 당해 차수 조사에 참여할 수 없는 조사 불가능 패널

* 조사불가지역 → 이사, 시설에 잠시 입소 후 복귀하는 경우 들이 있어 소멸표본으로 정의하긴 어려움

◦ 조사가능: 6,121명(원표본)을 기준으로 당해연도 조사불가를 제외한 나머지 패널을 의미한다.

◦ 이탈 후 진입 표본: i 차수조사 미참여 후, t차 조사 참여한 표본이다. $i < t,\ i = 2, \ldots, t-1$.

3 표본유지율 및 조사완료 현황

1) 표본유지율

장애인삶 패널조사의 3차년도 조사의 조사대상은 1차 년도 구축된 장애인 6,121명과 그들과 함께 거주하는 가구원이다. 2차년도 조사는 5,527명 장애인 패널을 조사 완료하였으며, 3차조사는 원표본 기준 5,259명 조사완료 하여, 약 85.9%의 표본유지율(=3차년도 조사완료자/1차 년도 패널구축 6,121명)을 보였다. 이는 타 패널조사의 표본유지율과 비교해도 높은 편이다.[17)]

〈표 Ⅱ-3-1〉 조사완료 현황

조사차수	조사완료 현황(명)			표본 유지율(%)	표본이탈	
	참여	미참여	계		(1차)응답 → 조사미참여(무응답)	조사불가
1차(2018)	6,121	-	6,121	100	-	-
2차(2019)	5,527	594	6,121	90.3	471	123
3차(2020)	5,259주)	862	6,121	85.9	572	290

주) 원표본 유지율(%) = (5,259/6,121)×100

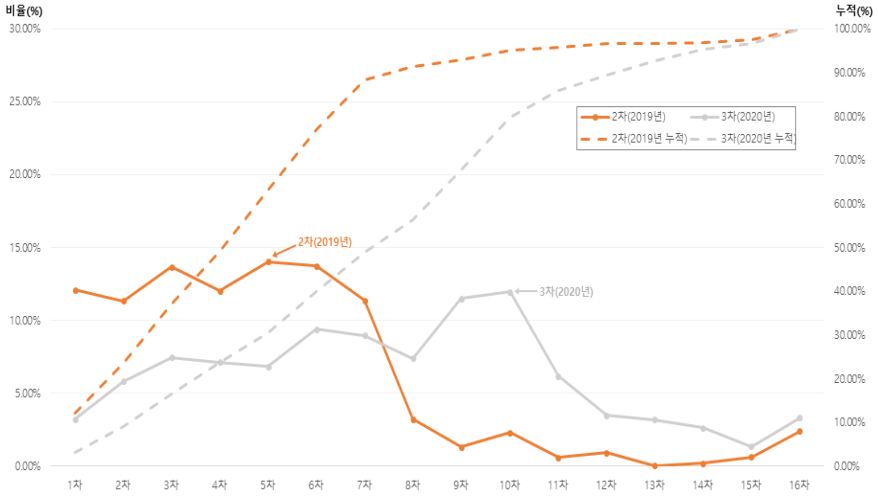
17) 〈표〉 타 패널조사 초기 표본유지율

조사차수	장애인삶 패널조사	한국노동패널	여성가족패널	한국청소년패널	
				중2	초4
1차	100.0	100.0	100.0	100.0	100.0
2차	90.3	87.6	84.5	92.4	95.2
3차	85.9	80.9	80.1	90.6	94.0

주 1. 한국노동패널: 「방하남, 장지연, 황수경, 김기헌, 김지경, 박시내.(2002). 한국노통패널 기초분석보고서(Ⅲ). 한국노동연구원.(p.2)」을 참고하여 작성함.
2. 여성가족패널: 이택면, 김승연, 우원규, 이주영, 강석훈.(2009). 2009년 여성가족패널조사 사업보고서 및 제2차 기초분석 보고서. 한국여성정책연구원.(p.59)」을 참고하여 작성함.
「이택면, 주재선, 송치선, 강석훈.(2011). 2011년 여성가족패널조사 사업보고서 및 제3차 기초분석 보고서. 한국여성정책연구원.(p.58-59)」을 참고하여 작성함.
3. 한국청소년패널: 「이경상, 김기헌, 임희진, 오소연, 소용준, 김유미.(2007). 한국 청소년 패널조사(KYPS) Ⅴ-조사개요 보고서-. 한국청소년정책연구원.(p.24-25)」을 참고하여 작성함.
「이경상, 김기헌, 조주연, 김현희, 박영실.(2006). 한국 청소년 패널조사(KYPS) Ⅴ-조사개요 보고서-. 한국청소년정책연구원.(p.24-25)」을 참고하여 작성함.

2) 연차별 조사완료 현황

[그림 Ⅱ-3-1]은 2차-3차 조사의 주차별 조사완료 현황이다. 실선은 주차별 조사완료율을 의미하며, 점선은 누적완료율이다. 조사 중반부 이후 조사 조사완료 표본 비율이 급격하게 떨어지는 것으로 보인다. 즉 전체 완료 표본의 약 90%가 조사 초반에서 중반사이에 조사완료 되고 있으며, 조사 중반에서 후반 사이의 완료 표본은 10%에 불과하다고 볼 수 있다. 결국 조사 초반~중반 동안 컨택이 얼마나 적극적으로 이루어졌는지에 따라 전체 조사진행율을 좌우하며, 이는 결국 전체 표본유지율을 결정하는 것으로 판단되며, 조사 초반에 패널 컨택 및 참여유도를 위한 방안(인센티브 등) 검토가 필요할 것이다.



[그림 Ⅱ-3-1] 주차별 조사완료 표본(2차(2019년), 3차(2020년)

주) 1차 조사는 패널구축이 함께 진행되어, 2차-3차와 조사 진행 특성이 상이하다. 이에 주차별 조사완료 현황을 대상으로 분석하였다.

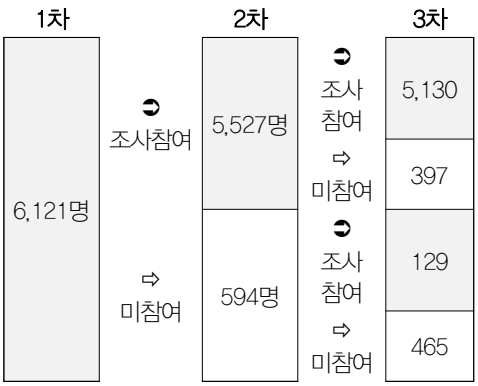
〈표 Ⅱ-3-2〉 연차별 조사참여 여부에 따른 응답패턴

구분	1차(2018)	2차(2019)	3차(2020)	계(명)
조사참여(명)	6,121	5,527	5,259	–
응답패턴1	1	1	1	5,130
응답패턴2	1	1	0	397
응답패턴3	1	0	1	129
응답패턴4	1	0	0	465
계				6,121

* 1: 조사참여, 0: 조사미참여

조사 연차별 조사참여 및 미참여 현황은 〈표 Ⅱ-3-3〉와 같이 3차 조사 참여 결과를 기준으로 4가지 응답패턴으로 구분 할 수 있다. 〈표 Ⅱ-3-4〉를 살펴보면, 2차 조사 참여 후 3차 조사 미참여자는 397명으로, 2차 조사 참여자중 약 7%가 3차조사에 미참여한 것으로 나타났다. 3차 조사기준으로 1-3차 조사 연속참여자는 83.81%(5,130명)이며, 2차 조사 미참여 후 3차조사 참여는 129명으로 약 2% 수준이다. 이들 표본은 3차 조사에 재진입(re-entrance)한 패널로 〈표 Ⅱ-3-3〉에서 '응답패턴3'과 같다. 이는 패널 3차 조사에서 처음 나타나는 응답패턴이다.

〈표 Ⅱ-3-3〉 연차별 조사참여 및 미참여 현황



응답패턴	빈도(명)	비율(%)
111	5,130	83.81
110	397	6.49
101	129	2.11
100	465	7.60
1: 조사참여 0: 조사미참여	6,121	100.00

출처: 여성관리자패널조사. (2015). 여성관리자패널조사에서 발췌하여 장애인삶 패널조사에 맞게 재구성함

4 3차년도 무응답자 특성 분석

표본이탈은 사유별로 구분할 수 있는데, 조사거부, 비수신 등으로 당해 연도 조사가 가능하였으나 최종적으로 조사에 참여하지 않은 경우와 사망, 이민 등조사가 불가능한 사유로 조사에 참여하지 않은 경우이다. 전자의 경우 당해 연도 조사에는 참여하지 않았으나, 향후 조사에 참여할 가능성이 존재한다. 이와 달리 후자는 조사대상자가 아니거나, 향후 더 이상 조사에 참여 할 수 없는 표본으로 조사 불가능한 소멸 표본이다. 전자인 이탈 표본은 비소멸 표본으로 즉, 무응답 표본으로 정의된다(홍민기와 최효미, 2014)[18]. 이렇게 표본이탈은 사유에 따라 서로 매우 다른 특성을 지니기 때문에 이를 하나의 무응답으로 간주하는 것은 무리가 있다. 따라서 표본이탈 사유를 나누어서 고려하는 것이 필요하다.

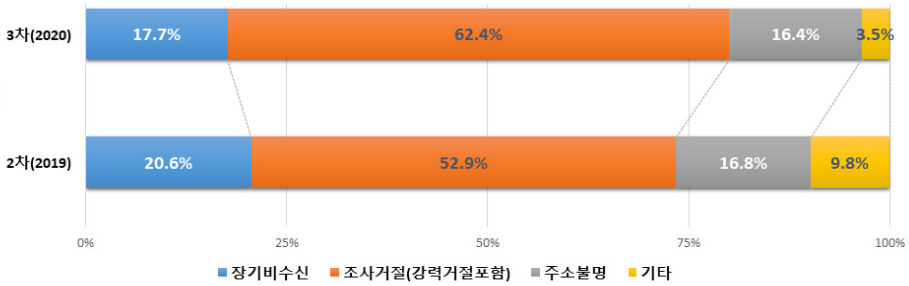
3차 조사에 미참여한 표본의 이탈유형을 살펴보면 조사 강력거절이 41.4%로 가장 높게 나타났다. 다음으로 사망이 23.5%, 장기비수신이 11.7%, 주소불명 10.9% 순으로 나타났다. 2차 조사의 이탈유형과 비교해보면 조사 강력거절이 가장 높게 나타난 것은 비슷하지만 2차 때는 장기비수신, 사망, 주소불명 순이었던데 비해 사망이 두 번째로 높게 나타났다는 차이를 보인다.

〈표 Ⅱ-4-1〉 조사차수별 조사참여 및 조사미참여 현황

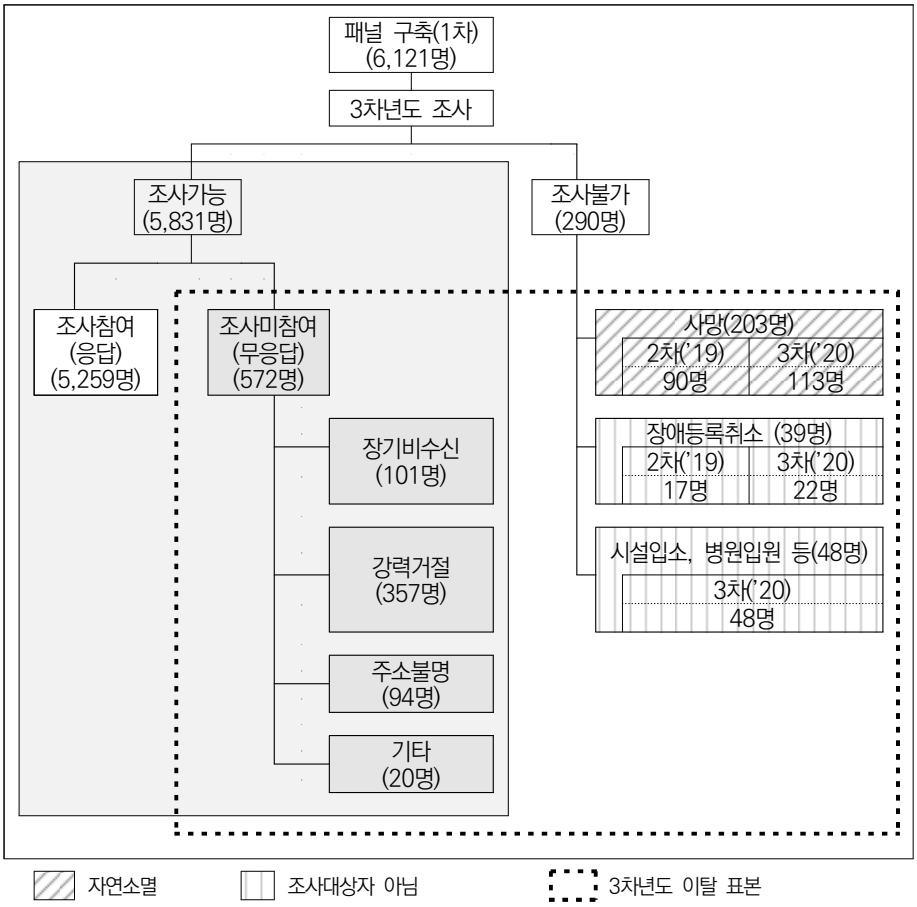
계		조사가능				소계	조사불가			소계	계
		장기 비수신	조사거절 (강력거절포함)	주소 불명	기 타		사망	시설입소, 병원입원 등	장애등록 취소		
2차(2019)		97	249	79	46	471	90	16	17	123	594
	(%)	20.6	52.9	16.8	9.8	100	73.2	13.0	13.8	100	–
	(%)	16.3	41.9	13.3	7.7	-	15.2	2.7	2.9	–	100
3차(2020)		101	357	94	20	572	203주)	48	39주)	290	862
	(%)	17.7	62.4	16.4	3.5	100	70.0	16.6	13.4	100	–
	(%)	11.7	41.4	10.9	2.3	–	23.5	5.6	4.5	–	100

주) 이전 차수의 사망, 장애등록 취소 인원이 포함된 수치임

18) 본 문단은「김현지 외 (2020). 장애인삶 패널조사 종단가중치 작성 방안 연구. 한국장애인개발원」에서 재인용하였으며, 원문은「홍민기, 최효미. (2014). 노동패널 표본이탈과 불평등 지표. 조사연구. 15(4), 95-122」에서 발췌하여 내용을 재구성 하였다.



[그림 Ⅱ-4-1] 연도별 조사미참여(조사가능 中) 사유 비교



주) 김현지, 김태용, 이혜수. (2020).에서 발췌하여 3차년도에 맞게 재구성함.
출처: Hidiroglou et al. (1993). in Paul P. Biemer and Lars E. Lyberg. 2003. Introduction to Survey Quality, p.86.; 변종석, 이석훈, 정구현. (2013).에서 발췌하여 본 패널조사 내용으로 재구성함

[그림 Ⅱ-4-2] 장애인삶 패널조사 '조사가능' 및 '조사불가' 도표(3차조사)

조사차수별 조사참여 및 조사미참여 현황은 〈표 Ⅱ-4-1〉와 같다. 이를 조사가능과 조사불가 표본으로 나누어 이탈 사유별로 도식화 하면 [그림 Ⅱ-4-2]와 같다.

〈표 Ⅱ-4-2〉 3차 조사 기준 표본이탈(조사미참여) 사유별 현황 Ⅱ-세부현황

구분	1차년도 조사완료 표본	이탈 표본	표본 이탈 사유							3차년도 조사완료 표본
			조사미참여(무응답)				조사불가			
			장기 비수신1)	강력 거절2)	주소 불명	기타	사망	시설입소, 병원 입원 등	장애 등록 취소	
전국	6,121	862	101	357	94	20	203	48	39	5,259
지체	974	118	12	51	14	1	25	7	8	856
시각	757	103	13	45	11	2	19	11	2	654
청각	761	72	5	43	6	3	11	1	3	689
언어	380	47	9	21	4	.	6	1	6	333
지적	401	49	8	18	10	.	8	4	1	352
뇌병변	980	156	18	60	16	2	45	10	5	824
자폐성	166	20	3	14	1	1	.	1	.	146
정신	328	51	8	24	8	.	6	5	.	277
신장	580	96	12	36	12	4	26	3	3	484
심장	120	13	1	6	1	.	4	.	1	107
호흡기	163	42	3	12	2	2	20	1	2	121
간	206	39	3	12	4	4	14	.	2	167
안면	30	8	1	7	.	.	.	.	.	22
장루. 요루	148	29	2	4	2	1	17	3	.	119
뇌전증	127	19	3	4	3	.	2	1	6	108
중증	3,242	509	65	191	53	11	153	28	8	2,733
경증	2,879	353	36	166	41	9	50	20	31	2,526
남자	3,389	427	47	195	55	10	122	25	18	2,917
여자	2,732	390	54	162	39	10	81	23	21	2,342

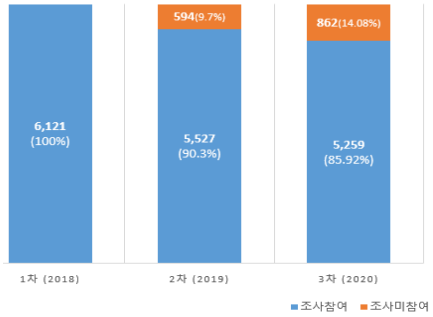
주 1) 장기출타 1명이 포함되어 있다.
2) 조사 중 강력항의로 인한 조사중단 1명이 포함되어 있다.

〈표 Ⅱ-4-3〉 조사차수별 조사참여 및 조사미참여 현황

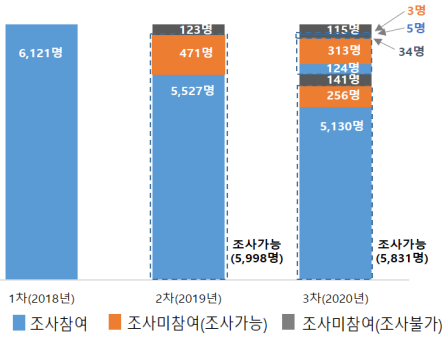
구분	1차 (2018)	2차 (2019)	3차 (2020)
유효표본주1) (유효표본 유지율)	6,121 (100%)	6,014 (91.9%)	5,879 (89.45%)
조사가능주2) (*사망, 장애등록취소, 시설입소·병원입원 등 제외)	6,121	5,998	5,831
조사참여 (표본유지율)	6,121 (100%)	5,527 (90.3%)	5,259 (85.92%)

구분			1차 (2018)	2차 (2019)	(%)	3차 (2020)	(%)	비고
조사미참여 인원			0	594	100	862	100	
조사미참여 사유	조사가능 -조사미참여 (무응답)	장기비수신	–	97	16.3	101	11.72	
		조사거절 (강력거절포함)	–	249	41.9	357	41.42	
		주소불명	–	79	13.3	94	10.90	
		기타	–	46	7.7	20	2.32	
		소계		471	79.3	572	66.4	
	조사불가	시설입소, 병원입소 등	–	16	2.7	48	5.57	
		사망	–	90	15.2	203주3)	23.55	3차 사망: 113명
		장애등록 취소	–	17	2.9	39주3)	4.52	3차: 22명
		소계		123	20.7	290	33.6	

주 1) 유효표본: 사망, 장애등록취소 제외를 제외한 인원
2) 조사가능: 당해연도 조사불가 패널(사망, 장애등록취소, 시설입소·병원입원 등) 제외
3) 이전 차수의 사망, 장애등록 취소 인원이 포함된 수치임



[그림 Ⅱ-4-3] 연차별 조사참여 및 미참여



주) 조사가능: 사망, 장애등록취소, 시설입소·병원입원 등을 제외한 표본

[그림 Ⅱ-4-4] 연차별 조사미참여 인원: 조사가능 및 조사불가

조사 차수가 경과하면서 조사불가 표본(사망, 장애등록취소 제외) 증가로 조사가능 표본 감소가 발생한다. '사망', '장애등록취소' 표본은 추적 조사 대상이 아니기 때문에, 지속적으로 증가할 것으로 예상된다. 따라서 조사불가가 아닌 '조사가능' 표본에 대한 관리가 필요할 것으로 보인다.

'조사가능' 패널의 조사미참여 사유를 살펴보면, 장기비수신, 주소불명에 비해 조사거절(강력거절포함)이 가장 높은 비중을 차지한다. '조사거절' 패널의 사유를 세분화 하고, 컨택 및 조사참여 유도를 위한 방안 강구할 필요가 있다. 조사미참여 사유가 조사거절인 패널이 2차조사에 비해 3차조사에서 약 9.5%p 높게 나타난다.

조사연차별 조사미참여 사유를 구체적으로 살펴보면, 강력거절을 포함하고 있는 조사거절 비중이 가장 높은 것으로 나타났다. 이러한 현상은 타 패널조사에서도 공통적으로 나타나고 있다. 각 연차의 조사가능 패널 중 조사거절 비율은 3차조사가 2차 조사에 비해 약 9.5%p 높게 나타는데 연차가 진행 될수록 조사거절 비율이 높아질 것으로 예상되는 부분이다. 조사참여 유도를 위한 방안에 대한 고민할 필요가 있다.

5 응답패턴 분석

본 절에서는 2차-3차 응답결과를 토대로 패널의 응답패턴을 분석하고, 응답패턴별 조사미참여 특성을 파악한다. 1차 조사는 패널구축이 동시에 이루어진바, 모든 조사대상이 조사에 참여하였다. 본 절에서는 조사대상 패널의 2차-3차 응답결과(〈표 Ⅱ-5-1〉)를 바탕으로 응답패턴별 특성을 분석한다.

〈표 Ⅱ-5-1〉 2차-3차 조사참여 현황

구분				3차 조사								계 (2차)
				조사가능					조사불가			
				조사참여	조사미참여				조사미참여			
					장기 비수신	조사거절 (강력거절 포함)	연락처 및 주소 불명	기타	사망	시설입소, 병원입원 등	장애등록취소	
2차조사	조사가능	조사참여		5,130	58	138	50	10	94	32	15	5,527
		조사미참여	장기 비수신	28	16	33	12	2	2	1	3	97
			조사거절 (강력거절 포함)	64	10	153	3	6	3	7	3	249
			연락처 및 주소 불명	18	12	22	20	1	6	.	.	79
			기타	14	5	9	8	1	6	2	1	46
	조사불가	조사미참여	사망	.	.	.	.	.	90	.	.	90
			시설입소, 병원입원 등	5	.	2	1	.	2	6	.	16
			장애등록 취소	.	.	.	.	.	.	.	17	17
계 (3차)				5,259	101	357	94	20	203	48	39	6,121

□ 2차 연속 응답: 5,130명

■ 2차 조사 이탈 후 3차조사 참여(재진입): 129명

▨ 2차 조사 응답 후 3차 조사 이탈 (조사불가(3차) 제외): 256명

〈표 Ⅱ-5-1〉의 2차-3차 조사참여 현황을 살펴보면, 2차년도 조사 불가 사유가 '장기비수신', '연락처 및 주소 불명'이었으나 3차년도 조사에 참여한 의 경우가 2차 조사가능 하였으나 미참여한 패널(471명)의 약 9.8%인 것으로 나타났다. 따라서 '장기비수신', '연락처 및 주소 불명'이 지속적인 패널 이탈로 이어지지 않을 수 있는바, 당해차수 조사에서 '장기비수신', '연락처 및 주소 불명'의 사유로 이탈하는 패널이 발생하지 않도록 유인할 수 있는 방안을 마련하는 것이 필요할 것이다.

3차 조사기준으로 응답패턴은 다음과 같이 네 개로 분류할 수 있다(〈표 Ⅱ-5-2〉). '1'은 해당차수에 조사참여, '0'은 조사미참여를 의미한다. 응답패턴1은 2차 조사 참여한 5,527명 중에서 3차 조사에 참여한 패널 이다. 이들은 1차-3차 조사[19]에 모두 참여한 장애인 패널로 5,130명이다.

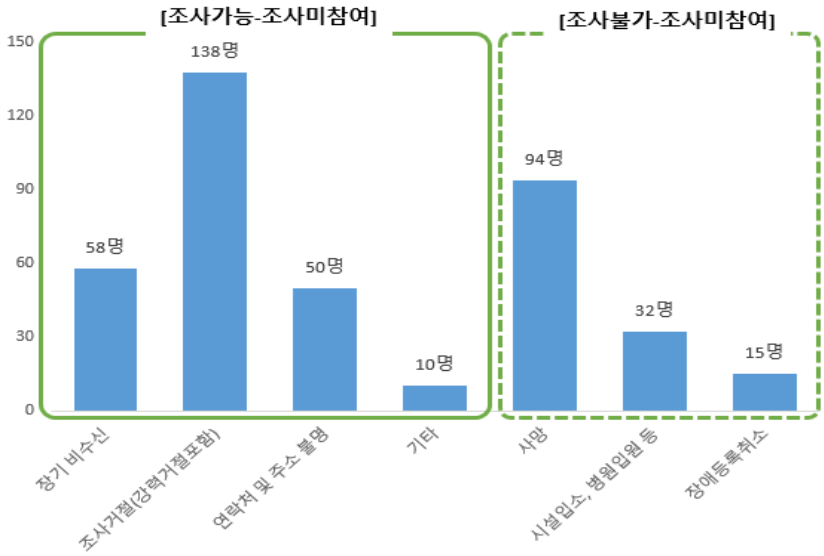
〈표 Ⅱ-5-2〉 응답패턴

구분	1차(2018)	2차(2019)	3차(2020)	계(명)
응답패턴1	1	1	1	5,130
응답패턴2	1	1	0	397
응답패턴3	1	0	1	129
응답패턴4	1	0	0	465
계				6,121

주) 1: 조사참여, 0: 조사미참여

응답패턴2를 가지는 패널은 397명으로 2차 조사 참여 후 3차조사에 미참여한 경우이다. 이들 중 141명은 사망, 장애등록 취소, 시설입소·병원입원 등으로 당해연도 조사가 불가능한 패널 이다([그림 Ⅱ-5-1]). [그림 Ⅱ-5-1]를 살펴보면, 조사가능하였으나 조사미참여한 256명 중 조사거절이 138명(53.9%)으로 가장 높은 비중을 차지하고 있으며, 연락처 및 주소 불명과 장기비수신이 108명(20%)으로 그 다음으로 비중이 높음을 확인할 수 있다.

19) 장애인삶 패널조사는 1차년도 조사와 패널구축이 함께 진행되었다.



[그림 Ⅱ-5-1] 응답패턴2(397명)-3차 조사 미참여 사유

응답패턴3은 2차조사 미참여 후 3차 조사 참여한 패널(129명)이다.

〈표 Ⅱ-5-3〉 응답패턴3

구분	1차(2018)	2차(2019)	3차(2020)	계(명)
응답패턴1	1	1	1	5,130
응답패턴2	1	1	0	397
응답패턴3	1	0	1	129
응답패턴4	1	0	0	465
계				6,121

〈표 Ⅱ-5-3〉에서 살펴보면, 2차조사 이탈 후 진입 표본(2차 조사 이탈 후 3차 조사 참여) 129명이다. 이들을 〈표 Ⅱ-5-1〉에서 2차-3차 조사 참여 결과로 분석해보면, 2차 조사에서 조사거절 하였으나 3차 조사에 참여한 패널은 64명으로 2차 조사 거절자(249명) 중 약 30%(=64명/249명)가 다음 차수 조사에 참여했다는 사실을 확인할 수 있다.

이들의(129명) 이전 이탈사유를 보면, 즉 49%(=64명/129명) 이상이 조사거절(강력거절 포함)로 조사미참여 하였으며, 21%이상(=28명/129명)이 장기비수신임을 확인할 수 있다. 다시 말해 3차 조사참여 결과를 기준으로 3차조사에는 참여하였으나, 2차 조사에서는 조사거절(강력거절 포함) 이력이 있는 패널이 49%이상(=64명/129명)이다. 종합해 보면, 강력거절의 경우 다음차수에도 조사 거절할 가능성이 높지만, 반드시 다음차수의 강력거절로 이어지지는 않는 다는 것이다. 따라서 당해차수 조사 미참여 사유가 강력거절인 경우, 구체적인 사유를 파라데이터로 수집하고 재분류하여 패널을 관리할 필요가 있을 것이다. 또한 2차년도 연락처 및 주소 불명으로 조사미참여 하였으나, 3차년도에 조사가능한 표본은 20% 이상(=18명/79명)인 것으로 나타났다. 이에 연락처 및 주소 변경을 패널이 알려올 수 있도록 유인책을 마련하는 것 역시 패널이탈을 방지하고 관리하기 위한 하나의 방법이라고 할 수 있다.

6 소결

패널조사는 동일 대상에 대해 매년 추적조사 하는 연구 방법으로 시간 경과에 따라 패널이 조사에 참여하면서 느끼는 피로도가 높아지면서 이는 궁극적으로 이탈율이 높아진다는 특징이 있다. 이는 국내·외 대부분의 패널에서 나타나는 현상이다(주재선 외, 2012; 이상호, 2013). 따라서 패널조사에서 탈락하는 개인 또는 가구의 특성을 파악하는 것은 무엇보다 중요한 요소 중 하나이다. 패널 이탈이 표본의 대표성 훼손 및 패널조사 결과를 활용한 추정치의 편의를 유발할 수 있기 때문이다. 이에 가중치를 통한 통계적 접근과 함께 표본이탈의 과정과 패턴에 대한 연구가 필수적이며(이경희와 민인식, 2015), 이를 토대로 이탈가능성을 예측하고 이탈예상 패널에 맞는 패널관리가 필요하다.

장애인삶 패널조사는 1차 년도 구축한 6,121명이 3차년도 조사대상이다. 본 조사의 2차년도 조사에서 부터 패널조사의 특수한 형태 무응답인 단위무응답 즉, 표본이탈이 발생하였다. 3차년도 조사에서는 2차 조사 미참여 후 3차 조사에 참여한 표본의 응답패턴인 재진입 패널이 나타났다. 국내·외 많은 패널조사

에서는 2차-3차 조사에서 표본이탈이 집중적으로 나타나는 것으로 알려져 있다.[20] 장애인삶 패널조사에서는 동일한 현상이 나타나면서 패널 관리의 중요성이 부각되고 있다. 이에 본 장에서는 1차-3차년도 조사참여 결과를 토대로 패널의 응답형태를 정의하고, 3차년도 무응답 특성 및 응답패턴을 분석하여 패널관리에 필요한 기초자료를 마련하고자 하였다.

분석결과를 종합해 보면, 당해 차수 조사거절로 조사에 미참여한 경우라도, 다음 차수 조사 참여를 유도할 수 있는 장치가 필요하다. 실제로 3차 조사에서는 재진입(re-entrant), 즉 2차 조사 미참여 후 3차 조사에 참여한 표본이 129명인 것으로 나타났다. 이들 3차조사 재진입 패널은 2차 조사 미참여자 (594명) 중에서는 약 22%(129명)에 해당하는 수치이다. 따라서 패널이탈 사유에 조사거절(강력거절 포함) 이력이 있다고 하여 향후 조사에서도 지속적인 이탈 의사를 가지고 있다고 단정 지을 수는 없다고 판단된다. 즉 조사거절(강력거절 포함)을 조사에서 제외하기 보다는 이전 년도 조사거절 이력이 있는 경우 조사 중·후반부 컨택 보다는 조사초반 적극적으로 컨택하여 참여를 유도하는 방안 등을 고려해 볼 필요가 있을 것이다. 또한 이들의 조사참여 독려를 위해 조사원 인센티브를 통한 패널의 조사참여 유도 등의 유인 및 관리 방안을 마련할 필요가 있으며, 조사초반의 조사참여를 어떻게 이끌어 낼 것인가에 대한 지속적 고민이 필요하다.

---

20) 표본이탈이 조사 초기(2~3차 조사)에 집중되기 때문에, 표본이탈 문제는 표본크기 및 표본설계 방법에도 영향을 미치게 된다. (출처) 최효미, 이상호, 성재민, 김기민, 배기준. (2012). 패널자료의 품질 개선을 위한 조사기법 연구–컴퓨터를 이용한 조사(CAPI) 방법–. 한국노동연구원.; 노동패널조사는 조사가 완료되면 소정의 선물을 지급한다. 그 외에도 가구원의 생일, 결혼기념일, 회갑 등이 있을시 축하편지를 발송하고 수집된 자료가 학술 및 정책적으로 유용하게 사용되고 있음을 알려주는 소식지도 발송한다. 이러한 노력에도 불구하고 조사초기였던 2차년도와 3차년도의 '조사성공률'(retention rate)은 각각 88%와 81%로 표본이탈이 예상보다 높게 나타났다. (출처) 이상호. (2005). 한국노동패널의 표본이탈분석 – 가구소득을 중심으로 –. 노동리뷰, 11, 66–80.

한국장애인개발원
www.koddi.or.kr

Ⅲ. 선행자료 분석

1 개요

본 장에서는 재진입 패널이 최초로 나타나는 3차년도 종단가중치 산출을 위해 국내 · 외 패널조사의 가중치 산출 과정을 분석하고, 이를 통해 장애인삶 패널조사의 종단가중치(3차년도) 산출시 벤치마킹할 필요가 있는 사항들을 도출하고자 한다. 종단가중치 산출을 위해서는 가중치 산출 이전의 패널 구축 시점 부터의 다양한 정보, 예를 들어 표본설계 및 추출과정, 이전 차수의 가중치 산출과정 등에 대한 이해가 선행되어야 한다. 이러한 사항들은 해당 조사의 분석보고서 등을 참고하였으나, 공개된 보고서만으로 종단가중치 작성 과정 전반에 대한 구체적인 내용을 파악하는 것에는 한계가 있다. 이에 일정 부분은 연구진이 추론하여 종단가중치 산출 과정을 정리하였다. 선행자료 분석을 토대로 3차년도 이후의 장애인삶 패널조사의 종단가중치 산출의 방향과 작성 방안을 마련하기 위한 기틀을 잡고자 한다.

다만 앞서 논의하였듯이 본 절의 선행자료 분석은 3차년도 이후의 종단가중치 산출에 대한 시사점 및 벤치마킹 사항을 도출하기 위한 것이다. 따라서 본 장에서 언급하는 조사들의 조사목적, 조사배경, 조사 전반에 대한 내용을 파악하는 것은 연구목적에서 벗어나는바, 본 장에서 이러한 사항들은 구체적으로 언급하지 않음을 사전에 밝혀둔다. 각 조사와 관련한 자세한 사항은 해당 조사의 연구보고서 등을 참고한다. 따라서 본 장에서의 검토 범위는 표본추출 및 전반적 가중치 작성 과정으로 제한한다.

2 타 패널조사 분석

1) 산재보험패널조사

(1) 3차년 이후 종단가중치 조정 과정[21)]

산재보험패널조사의 3차년도 가중치는 2차년도와 동일한 방식으로 산출되었다. 다만 3차년도 조사에 참여한 패널은 이전 차수 조사와 비교할 때 참여 특성이 상이한데, 1차 조사부터 3차 조사까지 모두 조사에 참여한 연속참여패널과 2차 조사 미참여한 후 3차 조사에 참여한 재참여패널이 나타난다는 것이다. 이러한 패널들의 조사참여 특징이 상이하여, 산재보험패널조사에서는 연속참여패널과 1회 이탈 후 조사에 참여한 패널로 나누어 두 가지 유형의 가중치를 산출하였다.

연속참여패널은 1,688명로 2차년도 가중치를 원가중치로 활용하였다. 2차년도 조사에 미참여한 후 3차년도 조사에 참여한 패널(36명)에 대해서는 권역, 장해유형, 재활서비스 이용여부, 성별를 활용하여 1차 년도 가중값 정보를 원가중치로 동일 그룹의 2차년도 종단가중치 평균으로 대체하여 적용하였다. 연속참여패널 대상의 Type1 종단가중치(변수명: wt031)는 시계열분석에 적합한 3차년도 종단가중치이며, 재참여패널을 포함한 3차년도 전체 참여 패널을 대상으로 산출된 Type2 종단가중치(변수명: wt032)는 3차년도 기초통계분석에 적합한 가중치이다.[22)] 이렇게 산출한 가중합은 Type1가 82,972명, Type2가 85,339명으로, 모집단 82,493명과는 각각 479명 및 2,846명의 차이를 보인다. 이는 층화변수인 권역 · 장해등급 · 재활서비스 이용여부별 모집단 수와 가중치 합계, 그리고 잠재 층화변수인 성 · 연령 · 요양기간별 모집단 수와 가중치의 합계에서도 차이가 발생한다. 이를 보정하기 위하여 산재보험패널조사는 2차년도와 같은 방식으로 사후층화(post-stratification)를 실시하였다. 사후층화는 레이킹 비 조정법을 적용하여 층화변수 교차셀과 잠재층화변수 교차셀별로 반복적 적용하여 가중값 보정한 값으로 최종가중값을 산출하였다.

---

21) 본 절은「이정화, 신슬비. (2017). 재4차 산재보험패널조사 기초분석 보고서. 근로복지공단 근로복지연구원.」에서 포괄적으로 발췌하여 요약 · 정리하였다.

22) 3차년도 종단가중값 산출을 위한 로지스틱 회귀분석 절차와 포함 변수들에 대한 설명은 2차년도 가중값 산출 내용을 참고하기 바란다.

(2) 장애인삶 패널조사에 주는 시사점

산재보험패널조사의 모집단은 2012년 1월~12월 사이에 요양 종결된 산재노동자에 대하여 시간 경과에 따른 변화양상을 분석하는데 의미를 두고 있다. 이는 장애인삶 패널조사의 모집단(2015.1.1 ~ 2017.12.31 사이에 장애등록을 마친 초기등록장애인)과 구조적으로 유사한 측면이 있다. 이에 본 패널조사 역시 구축 시점의 모집단이 조사 시점의 초기등록장애인을 대표한다고 해석하기엔 무리가 있다고 할 수 있다. 따라서 패널 구축 당시의 모집단이 시간 경과에 따라 어떻게 변하는 지를 분석할 수 있는 종단가중치 산출이 필요할 것이다. 이때 무응답 조정하여 산출한 가중치 합계와 모집단 총계간 차이는 레이킹 조정을 실시하는 것을 검토할 필요가 있다. 여기서 레이킹 보정을 위한 기준 모집단은 패널 구축 시점의 모집단을 활용한다. 산재보험패널조사는 3차년도부터 나타나는 연속참여패널과 이전 차수 이탈 후 재진입하는 패널을 특성을 반영하고 두 가지 유형의 가중치를 산출하고 있다. 장애인삶 패널 역시 연속참여패널과 재진입패널로 구분되는 특징이 3차년도 부터 나타나고 있는바 이러한 패널의 특성을 고려한 종단가중치 산출이 필요할 것이다.

〈표 Ⅲ-2-1〉 산재보험패널조사 종단가중치 현황

구분	레이블		조사참여자 수	가중치		모집단
				평균	합계	빈도
1차 코호트	1차 가중치		2,000	41.25	82,493	82,493
	2차 가중치		1803	45.75	82,493	82,493
	3차	3차 종단가중치(시계열분석용)	1,668	49.46	82,493	82,493
		3차 종단가중치(기초통계분석용)	1,704	48.41	82,493	82,493
	4차	4차 종단가중치(시계열분석용)	1,588	51.95	82,493	82,493
		4차 종단가중치(기초통계분석용)	1,660	49.69	82,493	82,493
	5차	5차 종단가중치(시계열분석용)	1,514	54.49	82,493	82,493
		5차 종단가중치(기초통계분석용)	1,616	51.05	82,493	82,493
2차 코호트	1차 가중치		3,294	51.29	75,392	75,392
	2차 가중치		2,965	59.78	75,392	75,392
	3차	3차 종단가중치(시계열분석용)	2,759	65.83	75,392	75,392
		3차 종단가중치(기초통계분석용)	2,823	63.37	75,392	75,392

2) 한국청소년패널조사(KYPS)[23]

(1) 3차년 이후 종단가중치 조정 과정[24]

3차년도 종단가중치는 1차-3차년도에 모두 응답한 패널(중2 패널 3,017명, 초4 패널 2,607명)만을 대상으로 산출한다[25]. 가중치는 두 단계를 통해 산출되는데, 첫 번째 단계는 응답확률 보정이다. 이는 로지스틱 회귀모형을 통해 응답확률을 추정한다. 모형에 투입되는 데이터는 2차 년도 조사를 통해 수집된 데이터 이며, 종속변수는 3차년도 응답여부를 이항변수로 정의하여 활용한다. 이때 '1'은 응답, '0'은 무응답을 의미한다. 로지스틱 회귀모형을 적용[26]하여 응답확률(propensity score)을 산출하고, 2차년도 종단가중치에 산출한 예측치의 역수를 곱하여 3차년도 종단가중치 산출을 위한 기본가중치를 계산한다.

$$\alpha^{l}_{s,hijk} = w^{l}_{s-1,hijk}\, p^{(l)}_{s-1,hijk}{}^{-1},$$

여기서, $\alpha^{l}_{s,hijk}$은 s차년도 종단면 기본가중치, $w^{l}_{s-1,hijk}$은 s-1차년도 종단면 가중치, 그리고 $p^{(l)}_{s-1,hijk}$은 응답확률 예측값을 의미한다. 최종적으로 종단 가중치는 모집단 변화를 반영하기 위해 해 사후층화하여 산출한다.

---

23) 한국청소년패널조사(KYPS: Korea Youth Panel Survey)의 조사 대상은 전국의 중학교 2학년 학생 중 표본으로 선출된 청소년을 대상으로 2003년(중학교 2학년)부터 2008년(고등학교 3학년 졸업 이후 1년 차)까지 6년 동안 반복적으로 추적조사하고, 전국의 초등학교 4학년 학생 중 표본으로 선출된 청소년 대상으로 2004년(초등학교 4학년)부터 2008년(중학교 2학년)까지 5년 동안 반복적으로 추적 조사하였다.

24) 본 절은 「박민규, 이경상, 박현수, 강현철. (2011). 한국청소년패널조사(KYPS) 가중치 부여 방법 연구: 중학교 2학년 패널의 경우. 조사연구. 12(3), 173-186.」과 「이경상, 안선영. (2009). 한국청소년패널조사Ⅶ : 1-6차년도 조사개요보고서」에서 포괄적으로 발췌하여 요약 · 정리하였다.

25) 패널이 전입/전출, 전학 등으로 이동이 있는 경우 표본으로 계속 유지되어 조사되며, 이전 무응답에 대한 원인 등을 파악하기 위한 별도의 추적조사가 실시되지 않는다.

26) 로지스틱 회귀모형의 적용에 앞서 주요 설명변수의 응답률 예측에 대한 효율성을 알아보기 위해 χ^2-검정을 통해 각 설명변수에 대한 3차년도 응답자와 무응답자 간의 차이를 비교하고, Wald의 적합도 검정을 바탕으로 한 단계적 방법(stepwise method)을 사용하여 로지스틱 회귀모형의 변수선택 방법을 고려하였다.

$$w^{l}_{s,hijk} = \alpha^{l}_{s,hijk}\left(\frac{N_{s,p}}{\widehat{N}_{s,p}}\right)$$

$w^{l}_{s,hijk}$: s차년도 종단면 최종 가중치
$\alpha^{l}_{s,hijk}$: s차년도 종단면 기본가중치
$N_{s,p} = \sum_{U} z_{s,p,hijk}$: 지역별/성별 모집단 크기
$\widehat{N}_{s,p} = \sum_{S} \alpha^{l}_{s,hijk} z_{s,p,hijk}$: 지역별/성별 기본가중치의 합계

* $z_{s,p,hijk}$ 는 h번째 층 내 i번째 학교의 j번째 학급에 속한 k 번째 패널이 s차년도에 사후 층 p에 속하면 1, 그렇지 않은 경우 0의 값을 갖는 지시 변수임

(2) 장애인삶 패널조사에 주는 시사점

한국청소년패널조사는 응답확률 보정시 로지스틱 회귀모형을 적용하여 응답확률을 예측하였다. 로지스틱 회귀모형의 종속변수는 3차년도 응답여부를 이항변수('1'은 응답, '0'은 무응답)로 정의하였으며, 설명변수는 3차년도 기준 응답자와 무응답자간 차이 분석 후 설정하였다. 이때 설명변수는 각 연차의 데이터를 활용하여 로지스틱 회귀모형의 변수선택 방법을 적용하여 변수를 선택하였다. 모형으로부터 얻어진 응답확률의 역수에 직전년도 가중치를 곱하여 개별 학생에 대한 s차 년도 종단가중치를 산출하였다.

장애인삶 패널조사 역시 3차년도 가중치 산출시 이전 년도와의 일관성 등의 측면을 고려하여 응답확률 추정을 위한 모형으로 로지스틱 회귀모형을 고려하고 있다. 이때 추정 모형을 위한 투입변수 선정을 위해서는 다양한 방법이 존재한다. 한국청소년패널조사처럼 각 연차별 영향을 미치는 변수를 선택할 수도 있고, 개념상 2차 조사의 응답지속여부와 3차 조사에서의 응답지속여부의 결정요인이 상이할 수 있다는 점을 고려하여 조사시점별 응답모형을 적용하는 것이 타당할 수도 있다(주재선 외, 2012).[27] 이에 투입변수 선정을 위해 조사참여 결과를 토대로 다양한 분석을 통해, 최종적으로 투입할 변수를 선정해야 할 것이다.

27) 주재선 외. (2012). 여성가족패널조사. 한국여성정책연구원.

3) 여성가족패널조사[28)]

(1) 3차년 이후 종단가중치 조정 과정

여성가족패널조사는 3차년도 종단가중치 산출[29)]을 위해서 2차년도 응답자를 기준으로 2차년도와 3차년도에 모두 응답한 개인과 3차년도에 응답하지 않은 개인에 대해 응답확률을 추정하는 로짓 모형을 적용하였다. 모형을 통해 추정한 응답확률의 역수를 2차년도 종단가중치에 곱하여 기본가중치를 산출하였다. 그리고 2차년도 2차년도 사후층화 정보(지역별, 연령별 적격응답여성수)를 이용하여 사후층화를 실시하여 최종 종단가중치를 산출하였다.

산출 과정에 대해 구체적으로 살펴보면, 3차년도에 응답한 개인 중 id를 골라 2차년도 가구원 파일의 long form에 머지하고 이를 통해 응답확률을 계산한다. 이때 활용한 변수는 연령, 연령제곱, 교육수준, 가구주여부(본인), 가구주여부(배우자), 결혼여부, 직장여부, 지역, 주거형태이다. 이들을 이용하여 추정한 응답확률모형 추정치는 다음과 같다. 산출한 응답확률의 역수를 2차년도 종단 가중치에 곱하여 사후층화전 가중치로 작성하고, 사망자를 조정한 후 최종적으로 3차년도 응답자 종단 가중치를 산출한다.

〈표 Ⅲ-2-2〉 2차년도 응답자 중 3차년도 응답자 응답확률 추정결과-여성가족패널조사

parameter	변수설명	자유도	추정값	표준오차	wald 검정	
					Chi-Square	Pr 〉 ChiSq
Intercept	-	1	2.2765	0.5525	16.9767	〈.0001
SAGE	연령	1	0.000606	0.000231	6.862	0.0088
AGE	연령제곱	1	-0.0752	0.0228	10.8637	0.001
edu1	교육수준더미	1	0.7217	0.2993	5.8127	0.0159
edu2	교육수준더미	1	0.7415	0.1923	14.8735	0.0001

28) 본 절은「주재선, 김영택, 송치선, 손창균, 임찬수. (2012). 2012년 여성가족패널조사. 한국여성정책연구원.」과 「김현지, 김태용, 이혜수. (2020). 장애인삶 패널조사를 위한 종단가중치 작성 방안 연구. 한국장애인개발원.」에서 포괄적으로 발췌하여 요약 · 정리하였다.

29) 여성가족패널의 종단가중치는 기록개인 종단가중치(가구에 존재하는 적격여성가구원을 분석 단위로 할 때 사용할 수 있는 가중치)와 응답자 종단가중치(가구에서 조사에 응답한 적격여성가구원을 분석할 때 사용되는 가중치)로 나뉜다. 장애인삶 패널조사의 경우 조사에 참여한 경우만을 대상으로 응답률을 추정하고 가중치를 부여하는 바, 응답자 종단가중치에 대해서만 검토하였다.

parameter	변수설명	자유도	추정값	표준오차	wald 검정	
					Chi-Square	Pr 〉 ChiSq
edu3	교육수준더미	1	0.6064	0.1797	11.3815	0.0007
edu4	교육수준더미	1	0.3879	0.1112	12.1644	0.0005
edu5	교육수준더미	0	0	.	.	.
rel1	가구주(본인)	1	0.7634	0.1667	20.9734	〈.0001
rel2	가구주(배우자)	1	-0.0387	0.1659	0.0546	0.8153
wed1	결혼여부	1	-0.8396	0.1792	21.9409	〈.0001
job1	직장여부	1	0.4312	0.0923	21.7982	〈.0001
DREG1	지역더미	1	-0.7886	0.2248	12.3019	0.0005
DREG2	지역더미	1	-0.6166	0.2605	5.6049	0.0179
DREG3	지역더미	1	-0.6106	0.2635	5.3698	0.0205
DREG4	지역더미	1	-1.3412	0.2981	20.2414	〈.0001
DREG5	지역더미	1	-0.8872	0.2744	10.4552	0.0012
DREG6	지역더미	1	-0.4925	0.2597	3.5977	0.0579
DREG7	지역더미	1	0.2546	0.3064	0.6905	0.406
DREG8	지역더미	1	-0.7035	0.2302	9.3352	0.0022
DREG9	지역더미	1	-0.3093	0.2589	1.4277	0.2321
DREG10	지역더미	1	-0.4783	0.2697	3.1468	0.0761
DREG11	지역더미	1	-0.0417	0.2907	0.0206	0.8859
DREG12	지역더미	1	-0.501	0.2679	3.4973	0.0615
DREG13	지역더미	1	-0.3231	0.2665	1.4692	0.2255
DREG14	지역더미	1	-0.4578	0.2504	3.3415	0.0676
DREG15	지역더미	1	-0.093	0.2572	0.1307	0.7177
DREG16	지역더미	0	0	.	.	.
dhten1	자가거주	1	18.3342	109	0.0283	0.8664
dhten2	전세거주	1	18.4052	236.7	0.006	0.938

주) 주재선, 김영택, 송치선, 손창균, 임찬수. (2012). 2012년 여성가족패널조사. 한국여성정책연구원.

여성가족패널조사는 3차년도 가중치 작성을 위한 고려사항으로 응답확률 모형, 모집단 층화변수, 가중치 절삭, 웨이부 무응답을 선정하여 각각의 검토를 진행하였다. 응답확률 추정모형은 로짓모형이 추정의 용이함 등의 장점이 존재하지만 로지스틱분포의 제약이 있음을 언급하였다. 현실적으로는 로지스틱 추정모형을 종단가중치 산출시 적용하였다.[30]

두 번째 고려사항은 응답확률 추정 모형 설정을 위해 투입할 설명변수이다. 이론적으로는 응답률의 차이가 크면서 변수간의 상관관계의 크기가 작고, 결측치가 없는 변수를 선정할 필요가 있다. 다만 모든 조건을 만족하는 변수를 실제 가중치 설계 단계에서 활용하는 것은 어렵다. 이에 항목무응답 보정 등의 방안의 검토 필요성을 언급하였다.

여성가족패널은 2차년도까지 사후층화에 사용한 모집단 정보는 조사과정에서 나타난 정보를 모집단 정보와 결합하여 층화한 유사모집단정보이다. 이는 여성가족패널의 적격응답개인은 19~64세 여성이지만 추출단위가 가구인 관계로 모집단의 조율이 복잡하고, 활용할 수 있는 모집단 정보가 없기 때문이다[31]. 이에 3차년도 가중치 작업을 하는 경우 2010년 센서스자료의 사용가능 여부의 검토 필요성을 검토하였다. 가중치 절삭에 대해서도 검토가 진행되었는데, 이는 응답확률이 매우 낮게 추정되어 개인의 가중치가 매우 크게 나타나는 경우를 대비하여 가중치의 분산을 적정수준으로 만들 필요가 있기 때문이다[32]. 일반적으로 표준편차의 2배 또는 3배의 방식으로 절삭하기도 하며[33], 통계적 기법을 이용하여 추정량의 편향(bias)과 분산의 합을 최소화하는 방법으로 가중치를 조정하기도 한다.

여성가족패널조사 역시 타 패널조사와 마찬가지로 3차년도 조사부터 재진입패널, 연속응답패널이 발생하였다. 즉 2차년도 응답하지 않았으나, 3차년도 응답한 패널과, 모든 조사 차수에 응답한 패널 이다. 이들은 종단가중치의 대상이며, 2차 조사 참여 후 3차 조사에 참여하지 않은 패널은 3차년도 기준의 종단가중치 대상이 아니다. 이 외 에도 응답확률을 추정할 때 사망이나, 해외이민, 집단시설 수

30) 다만, 새로운 모형을 도입시 기존에 작성한 가중치를 전부 새롭게 계산해야 한다.
31) 연구진은 2005년 센서서를 모집단 정보로 사용할 수 있으나, 시간차이로 인해 의미가 없다고 판단하였다.
32) 이러한 문제는 이론상으론 추정량의 불편성을 유지하는데 도움이 되지만, 가중치의 큰 분산은 추정량의 분산을 크게 하는 문제로 이어진다.
33) 개별가중치와 가중치의 평균의 차이가 표준편차의 2배 또는 3배가 넘는 경우, 개별가중치는 표준편차의 2배 또는 3배를 더하거나, 뺀 값으로 대체하는 방법이다.

용 등에 의해 응답하지 못한 경우는 응답확률 모형에서 응답한 것으로 간주하고 모형을 추정하고, 이 후 사후층화를 통해 이를 보정하였다.

(2) 장애인삶 패널조사에 주는 시사점

여성가족패널조사의 종단가중치는 2차년도 응답자를 기준으로 2차년도에 이어 3차년도에도 지속적으로 응답한 개인들과 3차년도에 응답하지 않은 개인들에 대해 응답확률을 추정하는 로짓 모형을 설정하고, 도출된 응답확률의 역수를 2차년도 종단 가중치에 곱하여 산출한다. 다음으로 2차년도 사후층화 정보(지역별, 연령별 적격응답여성수)를 이용하여 사후층화를 실시하여 최종종단가중치를 산출하였다.

장애인삶 패널조사 역시 2차년도 응답자를 토대로 3차년도 응답결과를 기준으로 하는 응답확률 추정 모형을 적용하는 것을 검토할 필요가 있다. 2차 조사에 참여한 후에 3차 조사에 참여한 대상자와, 2차 조사에서 1회 이탈 후 3차 조사에 참여한 패널의 특성이 상이할 것이기 때문이다. 응답확률 추정 모형 산출시 2차년도와 마찬가지로 응답 특성이 상이한 조사불가 패널을 제외할 필요가 있을 것이다. 응답확률 모형에서 활용할 설명변수도 중요한 고려 사항인바, 이론적 관점에서 응답률의 차이를 보이고 응답여부와 연관성이 있는 변수를 선정하기 위한 방안을 검토할 필요가 있다. 여성가족패널조사에서 검토한 측면들에 대해 본 패널조사의 종단가중치 산출시에도 검토할 필요가 있으며, 응답확률 추정시 사망 등 응답이 불가능한 패널을 응답확률 모형에서 어떻게 처리할지를 검토할 필요가 있을 것이다. 이는 2차년도 가중치 산출시 고려한 사항을 반영하는 것이 필요할 것이다. 또한 응답확률 추정 모형은 가중치 작성의 일관성 측면에서도 검토가 필요할 것이다.

[참고] 해외패널조사

1) PSID[34][35]

PSID의 종단가중치는 기본적으로 '역(逆)의 확률 가중치'로 2단계의 과정을 통해 작성된다. 첫 번째 단계에서는 조사에 참여한 모든 개인[36]에 대해 개인 종단가중치를 산출하고, 두 번째 단계에서는 조사에 참여한 모든 가족 구성원(표본 개인과 비표본 개인 모두 포함)의 개인 종단가중치를 평균하여 가족 종단가중치를 산출한다[37].

구체적으로 살펴보면, 개인 종단가중치를 구하기 위해서는 3단계의 과정을 따른다. 첫째, 해당 개인은 2차례의 연속적 조사(이전 조사(t_0), 현재 조사(t))에 대한 참여 성격(6가지 기준)에 따라 10개의 층위로 분류한다. 6개의 기준은 다음과 같다.

① 과거이탈(past exit): 과거 한 시점에 조사에 참여하였으나, 이전 조사(t_0) 이전에 이탈한 경우

34) PSID(Panel Study of Income Dynamics)은 소득과 빈곤의 역학 연구를 위해 미시간대학교(University of Michigan) 조사연구센터(The Survey Research Center, SRC)에서 실시한 최초의 패널조사로, 1968년 처음 조사를 시작한 이후 1997년까지는 1년주기로 조사하였고, 1999년 부터는 2년 주기로 조사를 수행하였다. 초기 PSID의 표본은 2개의 하위표본으로 구성되어 있으며, 표본설계 당시 목표 표본 수는 약 5,000가족이었다. 이 중 약 2,000가족은 경제기회조사(Survey of Economic Opportunity, SEO) 표본 중 저소득가구를 부표집(sub-sampling)하였다고, 나머지 약 3,000가족은 SRC에서 전국을 대표할 수 있도록 2단계 집락추출로 표본을 선정하였다. 초기에 구축한 표본은 1968년 이후 미국에 온 이민자를 대표하지 못한다는 한계가 있으며, 이에 1997년-1999년 조사와 2017년-2019년 조사에는 표본 내 이민자의 대표성을 확보하고자 각각 1997년 이민자 표본(1968년 이후 이민자)과 2017년 이민자 표본(1997년 이후 이민자)을 추가하였다.

35) 본 절은「정미옥, 백지선. (2011). 가계금융조사를 위한 가구패널 설계 연구. 통계청 통계개발원」,「김태완, 이병재, 이주미, 최준영, 김기원, 박수지, 송치호, 주찬희. (2018). 2018년 한국복지패널 심층분석 - 해외 주요 패널과의 비교 연구. 한국보건사회연구원.」및「Chang, W., Nishimura, R., Heeringa, S., Johnson, D., and Sastry N. (2021). panel study of income dynamics construction and evaluation of the 2019 longitudinal individual and family weights. survey research center, institute for social research university of michigan, Ann Arbor. 」에서 발췌하여 요약 · 정리하였다.

36) PSID에서는 개인을 표본 개인(sample individual)과 비표본 개인(nonsample individual)으로 분류한다. 표본 개인은 1968년 첫 번째 조사 당시 원표본 가족과 함께 살고 있던 개인과 첫 번째 조사 이후 태어난 그들의 직계비속이고, 비표본 개인은 표본 개인과의 결혼 등으로 표본으로 진입한 개인을 의미한다.

37) 장애인삶 패널조사의 조사대상은 장애인 패널과 그들과 함께 사는 가구원이나, 가중치는 패널에게만 부여한다. 이에 본 절에서는 PSID의 개인 종단가중치에 대해서만 검토하고자 한다.

② 미래참여(future entry): 현재까지 조사에 참여한 적은 없지만 미래의 한 시점에 참여할 경우
③ 최근참여(recent entry): 현재 조사(t)에 처음으로 참여한 경우
④ 재참여/재시도(reentry/recontact): 이전 조사(t_0)에 이탈하였으나, 현재 조사(t)에 다시 참여한 경우
⑤ 최근이탈(recent exit): 이전 조사(t_0)에 참여하였으나, 현재 조사(t)에 이탈한 경우
⑥ 응답(response): 이전 조사(t_0)와 현재 조사(t)에 모두 참여한 경우

6개 기준 중 최근이탈과 최근참여는 각각 2개, 5개의 층위로 구성하고, 그 외 나머지는 1개 층위로 구성한다. 이에 총 10개의 층위로 분류되며, 각 층별로 초기 가중치를 다음과 같이 부여한다.

현재조사에 참여하지 않은 경우는 가중치 0을 부여하고, 현재조사에 참여하였으나 비표본에 해당되는 경우 또한 0을 부여한다. 재참여/재시도에 해당하는 경우는 과거 마지막 조사 시기의 가중치와 동일한 값을 부여하고, 최근참여 중 원표본으로부터 출생/입양된 경우는 원표본가구의 가구주와 배우자의 가중치 평균값을 부여한다. 표본 추가 등으로 현재조사에 새롭게 참여한 표본의 경우는 해당 특성을 가진 개인에 대한 추출확률의 역수값을 부여한다.

둘째, 응답집단의 대표성을 유지하기 위해 시간의 경과에 따라 발생한 패널 마모(attrition)에 대한 조정을 한다[38]. 이에 이전 조사(t_0) 참여자가 현재 조사(t)에 참여할 수 있는 3가지 형태(응답, 사망, 무응답)와 생존여부(사망/생존)에 대한 조건부확률을 다항 로직스티 회귀분석을 통해 추정하고, 이 값을 통해 산출된 패널 마모 조정요인(attrition adjustment factor)[39]을 응답자 개인의 이전조사 가중치에 곱하여 현재조사 가중치를 산출한다.

위 내용을 토대로 정리한 초기가중치 부여 규칙은 다음과 같다.

38) 패널 마모의 처리에 대해 일반적으로 조정가중치(the adjusted weights)와 이월 가중치(the carry-over weights)를 통해 조정한다. 조정 가중치 방식은 표본 마모를 고려한 방식으로 개인의 이전 조사 가중치에 패널 마모 조정요인을 곱한다. 이월 가중치 방식은 표본 마모를 고려하지 않은 방식으로 개인의 이전 조사 가중치를 그대로 적용한다. PSID는 조정 가중치 방식을 적용하고 있다.

39) 마모 조정요인의 구체적인 산출방법은 「Gouskova E., S. Heeringa, K. McGonagle, R. Schoeni, and F. Stafford, (2008). Panel Study of Income Dynamics Revised Longitudinal Weights 1993-2005, PSID

〈표〉 PSID 초기 가중치 부여 규칙

구분		자료 수집 여부		초기 가중치
		이전조사(t_0)	현재조사(t)	
과거이탈 (past exit)		X	X	0
미래참여 (future entry)		X	X	0
최근참여 (recent entry)	원표본으로부터 출생/입양한 경우	X	O	원표본가구의 가구주와 배우자의 가중치 평균값 부여
	조사에 이탈한 개인에 의해 출생/입양한 경우	X	O	0
	표본 개인과 결혼/동거한 경우(비표본)	X	O	0
	새표본	X	O	해당 특성을 가진 개인에 대한 추출확률의 역수 부여
재참여/재시도 (reentry/recontact)		X	O	과거 마지막 조사 시기 가중치와 동일한 값 부여
최근이탈 (recent exit)	t_0와 t 사이에 사망한 경우	O	X	0
	t_0 응답 & t 무응답	O	X	0
응답 (response)		O	O	t_0 X 마모조정요인

주: 「정미옥, 백지선. (2011). 가계금융조사를 위한 가구패널 설계 연구. 통계청 통계개발원」〈표 5-4〉 재구성

셋째, 가중치를 표준화(scaling) 한다. 이는 개인 가중치를 1차 웨이브의 초기 가중치 총계에 맞춰 조정하는 단계로, 보정 방법으로는 ① 미국의 센서스 인구조사(Census)나 상시인구조사(Current Population Survey) 추정값을 벤치마킹하거나 ② 가중치를 표본크기에 맞게 조정하는 방법이 있다. 각각의 산식은 다음과 같다.

Technical Series Report #08-05, ISR, University of Michigan」 4.3 Attrition adjustment factor을 참고하기 바랍니다.

① $W_{i,pop} = W_i \cdot \frac{N_{pop}}{\sum_{i=1}^{n} W_i}$ ② $W_{i,norm} = W_i \cdot \frac{n}{\sum_{i=1}^{n} W_i}$

W_i : PSID 가중치

N_{pop} : 목표모집단인 센서스나 상시인구조사의 총합

n : 분석에 포함된 표본가구 또는 개인의 수

2) SOEP[40][41]

패널조사의 가중치는 일반적으로 설계가중치, 무응답 조정, 사후 조정의 3단계를 거쳐 산출된다. 패널조사의 모집단 총합을 추정하기 위해 불편추정량인 Horvitz-Thompson(HT) 추정량을 이용하며, 이를 위해 추출확률의 역수를 부여하는 방법을 활용한다. SOEP는 원가구 추출확률을 추정하기 위해 로짓모형을 이용한 방법을 적용하고 있다.

구체적으로 살펴보면, 1차 조사의 가구 가중치는 가구의 설계가중치에 무응답 조정과 벤치마킹 조정을 실시하여 산출하고, 가구 가중치를 모든 가구원의 가중치로 부여한다. 2차 조사 이후의 가구 가중치를 산출하기 위해서는 기존 가구원에게는 1차 조사 최종 가중치를 초기 가중치로 부여하고, 새로 진입한 가구원에게는 로짓모형을 이용하여 추정한 추출확률을 통해 초기 가중치를 부여한다. 초기 가중치에 무응답 조정과 벤치마킹 조정을 실시하여 최종 가구 가중치를 산출한다. 가구원의 가중치는 최종 가구 가중치를 모든 가구원의 가중치로 부여한 후, 무응답 조정과 벤치마킹 조정을 실시한다.

40) SOEP(Socio Economic Panel)는 독일 인구 및 소득 통계의 수립을 위해 독일 주요 경제연구기관인 DIW Berlin(Deutsch Institut für Wirtschaftsforschung Berlin)에서 실시한 유럽 최초의 패널조사로, 1984년부터 매해 고용, 교육, 자산, 소득, 개인이 가지고 있는 정부 정책에 대한 가치관 등을 조사하고 있다. SOEP는 조사 대상이 다양하여 표본의 구성이 복잡한 형태를 이루고 있다. 1984년 1차 조사 당시 표본은 2개로, 서독 지역에 거주하고 있는 가구주가 독일인인 4,528가구(표본 A)와 서독 지역에 거주하고 있는 가구주가 외국인인 1,393가구(표본 B)로 구성하였다.

41) 본 절은 (2018). 2018년 한국복지패널 심층분석 - 해외 주요 패널과의 비교 연구. 한국보건사회연구원.」 및 「Rainer Siegers, Veronika Belcheva, and Tobias Silbermann. 2020. SOEP–Core v35 - Documentation of Sample Sizes and Panel Attrition in the German Socio–Economic Panel (SOEP) (1984 until 2018). SOEP Survey Papers 826: Series C. Berlin: DIW/SOEP」에서 발췌하여 요약 · 정리하였다.

한국장애인개발원
www.koddi.or.kr

Ⅳ. 장애인삶패널 가중치 조정에 대한 고찰(3차년도)

1 기본방향

가중치는 조사단위가 추출확률, 무응답, 사후층화 등으로 서로 상이한 대표성을 가질 때 이를 보정하고자 조사단위별로 부여한다.[42] 종단가중치는 조사시점의 반복으로 인해 산출하는 가중치로, 일반적인 패널조사의 1차 조사에서는 횡단조사 가중치부여 방법을 그대로 사용할 수 있다. 그러나 종단가중치의 경우에는 패널조사에서 발생하는 특수한 형태의 무응답인 표본탈락(panel attrition)[43]과 비표본가구원의 진입을 고려해야 한다.[44] 종단가중치는 이전 차수의 가중치가 기본이며, 이전 차수의 응답데이터를 토대로 산출한다. 그러나 패널조사는 조사 차수가 진행되면서 표본 이탈이 발생한다는 특성이 있다. 이러한 특성으로 인해 이전 차수 가중치를 당해차수에 그대로 적용하는 것은 불가능하며, 표본 이탈 특성을 반영한 새로운 가중치 산출이 필요하다.[45]

장애인삶 패널조사의 경우 추출단위가 개인, 조사단위가 가구인 개인단위 패널조사 이다. 따라서 추적원칙에 비표본가구원의 진입은 포함되지 않는다. 이에 종단가중치 산출에서 비표본가구원의 진입은 고려 대상은 아니지만, 개인 패널의 표본탈락은 고려가 필요하다. 본 장에서는 패널의 표본 탈락을 고려한 장애인삶 패널조사의 3차년도 종단가중치 작성방안을 검토하고자 한다.

2장에서 논한바와 같이 장애인삶 패널조사에서의 3차년도 조사에서는 당해 차수의 표본 이탈 이외에도 2차년도 조사에서 이탈 한 후 3차년도에 재진입(re-entrance)하는 패널 유형이 나타난다. 즉 조사에 연속적으로 참여한 연속 참여 패널과 재진입 패널이다. 이들은 서로 상이한 특성을 가지기 때문에 이를 반영한 두 가지 유형의 가중값 산출이 필요하다. 하나는 연속 참여 패널에 부여된 가중치로 시계열분석에 적합한 3차년도 종단가중치이다. 다른 하나는 3차년도 응답한 모든 패널, 즉 재진입 패널을 포함하여 산출한 종단가중치이며 기초통계분석을 위한 것이다.

---

42) 김규성, 황영은, 박진우. 패널조사에서 가중치 부여 방법 및 효과에 관한 연구.
43) 이러한 전체 무응답 이외에도 개별 항목에 대해서만 응답을 거부하는 항목무응답(item nonresponse)도 발생할 수 있으나 이 경우는 가중치부여 방법 보다는 보정(imputation) 방법이 주로 사용된다. (출처) 강석훈. KLIPS 2차 웨이브의 가중치 부여방법에 관한 연구.
44) 강석훈. (2000). KLIPS 2차 웨이브의 가중치 부여방법에 관한 연구. 제2회 한국노동패널 학술대회 논문집.
45) 강석훈. (2003). KLIPS의 가중치 부여방안 연구. 한국노동패널연구 2003-4, 한국노동연구원.

기본적으로 장애인삶 패널조사 3차년도 종단가중치는 1차~3차 조사에 모두 참여하고, 3차 조사를 최종 완료한 5,259명에 부여한다. 기본가중치를 무응답 조정하여 기본가중치로 정의, 레이킹 비 조정하여 최종적으로 종단가중치를 산출한다. 최종적인 가중치 산출은 표본설계 및 연차별 가중치 작성방법의 일관성, 공표자료와 향후 데이터 활용성 측면을 전반적으로 검토하여 결정한다.

2 3차년도 종단가중치 작성의 고려사항

3차년도 조사에 참여하여 응답완료한 모든 패널을 대상으로 3차년도 가중치가 부여된다. 종단가중치는 이전 차수의 가중치를 기본으로 산출되는데, 장애인삶 패널조사는 〈표 Ⅳ-2-1〉의 응답패턴3[46]과 같은 이전 차수에 응답하지 않았으나, 당해 차수 응답한 재진입 패널이 존재한다. 이들은 이전 차수 가중치가 부재하여 종단가중치 산출시 고려할 이전 차수 가중치가 없다. 즉 응답패턴1과 같이 1차~3차 조사에 모두 참여한 연속응답패널과의 응답특성(무응답 특성)을 고려한 가중치 산출이 필요하다. 따라서 종단가중치 산출시에 3차년도 조사결과에 따른 응답패턴의 특성을 고려할 필요가 있다. 다시말해 가중치 산출시 고려가 필요한 첫 번째는 연속참여 패널의 가중치 산출은 어떤 방법을 적용할지에 대한 방안 검토이다. 둘째로는 종단가중치는 기본적으로 직전년도 가중치를 기반으로 산출되는데, 재진입 패널에 대한 가중치를 어떻게 처리할 것인지에 대한 것이다.

〈표 Ⅳ-2-1〉 가중치 부여 대상

구분	계(명)	1차(2018)	2차(2019)	3차(2020)		
				응답여부	가중치 부여 대상	비고
응답패턴1	5,130	1	1	1	O	
응답패턴2	397	1	1	0	–	
응답패턴3	129	1	0	1	O	2차 가중치 없음
응답패턴4	465	1	0	0	–	
계(명)	6,121	6,121	5,527	5,259		

46) 3차년도 응답결과에 따라 응답패턴1과 응답패턴3의 패널로 구분할 수 있다.

3 응답 여부와 주요 특성 변수들 간의 연관성

장애인삶 패널조사의 조사대상은 장애인 개인이다. 그리고 그들과 함께 거주하고 있는 가구원을 조사하고 있다. 종단가중치는 표본 이탈에 대응하기 위한 통계적 보정 방법이지만, 이러한 통계적 접근 뿐 아니라 연구 전반의 과정에서 패널 이탈을 최소화하기 위한 노력이 병행되는 것이 중요하다. 따라서 탈락하는 개인 특성 파악이 무엇보다 중요한 과제이다. 이를 위해 조사 결과를 바탕으로 응답 패널과 무응답 패널간 이탈 특성 분석을 실시하여 데이터 품질 관리를 위한 분석 기반을 마련하고자 한다.

본 절에서는 3차년도 무응답 패널과 응답 패널간의 차이를 특성 변수간 분석을 통해 파악한다. 그리고 이를 토대로 패널 이탈에 영향을 미치는 요인을 분석한다. 따라서 본 절에서는 3차년도 조사 결과를 토대로 조사참여에 여부를 종속변수로, 개인 및 가구, 금융특성과 관련한 변수에 대해 t-검정, χ^2-검정을 시시하여 차이를 비교하고, CHAID 분석를 실시하여, 응답여부와 주요 변수간의 연관성 분석을 실시한다. 조사참여에 영향을 미치는 특성 변수 파악 및 연관성 분석은 종단가중치 산출시 무응답 조정 모형의 기초정보로서 장애인삶 패널조사에 의미 있는 정보를 제공할 수 있을 것이다.

1) 분석대상

본 절에서는 무응답 조정시 고려할 특성 변수 분석은 기본적으로 1차 년도 조사 완료된(구축패널)[47] 6,121명을 바탕으로 실시한다. 종단가중치는 이전 차수의 응답 데이터를 토대로 산출되기 때문에, 1차-2차년도 조사에 참여한 5,527명의 3차년도 조사 참여 여부를 토대로 분석대상을 분류한다. 이 때 재진입[48] 패널은 이전 차수의 가중치가 존재하지 않기 때문에 재진입 패널은 분석에서 제외한다. 다시 말해 이전 차수의 가중치가 존재하지 않는 재진입 패널의 가중치 산출은 3차년도 이외에 1차-2차년도 조사에 모두 참여하여 이전 차수 가중치가 존재하는 패널의 가중치 산

47) 장애인삶 패널조사는 1차 년도 조사(2018년)와 패널구축이 동시에 이루어 졌다. 이에 최종적으로 6,121명의 장애인 표본을 패널로 구축하였다.

48) 앞서 논한바와 같이 3차년도 조사를 기준으로 3차 조사에는 참여하였지만 2차 조사에는 참여하지 않은 패널을 의미한다.

출 방법과는 상이한 바 이에 대해서는 다음절[49]에서 살펴보고자 한다. 본 절에서는 3차년도 기준 재진입 패널은 제외하고, 이전 차수의 응답데이터가 있는 2차년도 조사 완료 패널의 3차년도 조사 참여 결과를 바탕으로 분석대상을 설정한다.

장애인삶 패널조사 3차 조사에 참여하여 조사 완료한 패널은 5,259명이며 3차 조사 결과를 토대로 조사 참여와 미참여로 재분류하였다. 이 때 사망, 장애등록취소의 이유로 패널에서 영구 탈락되어 향후 차수부터는 조사대상이 아닌 패널, 그리고 당해 차수 조사 제외 대상인 시설입소 등은 분석에서 제외하였다. 따라서 최종 분석대상은 응답확률 추정 모형 도출시 투입하는 5,386명이다. 이들은 3차 조사 참여 결과를 토대로 다시 분류할 수 있는데, 조사에 참여한 경우는 조사에 협조적인 성향이 높은 패널로 볼 수 있을 것이다. 조사 미참여 패널은 사실상 비협조성향이 높은 패널로 간주하였다.

3차 가중치 처리를 위해 분석대상은 다음과 같다. 2차 조사 참여 완료한 패널을 대상으로 3차 조사 응답 결과를 기준으로 조사제외 사유가 사망, 장애등록취소, 시설입소 등인 패널은 제외하고 분석대상으로 설정한다. 이는 2차년도 가중치 산출시와 일관성을 유지하면서 조사 미참여 사유가 조사불가로 인한 것인지, 조사 가능하였으나 조사 미참여한 것인지에 따라 패널의 특성이 다를 것이라는 가정 하에서 진행되었다.

49) 제4절. 3차년도 가중치 처리 방안 비교 · 검토

〈표 Ⅳ-3-1〉 분석대상

구분	1차 조사 완료3)	1차 조사 완료	분석 2차 조사 완료
	6,121	6,121	5,527
	⬇	⬇	⬇
	2차 조사 결과	3차 조사 결과	3차 조사 결과
분석대상	5,998	5,831	5,386
조사참여(협조성향)	5,527	5,259	5,130
조사미참여(비협조성향)	471	572	256
분석제외			
사망1), 장애등록취소[1], 시설입소 등(조사제외)[2]	123	290	141

주: 1) 사망, 장애등록취소는 영구탈락
2) 시설입소는 일시적(당해 조사에서)으로 조사에 참여 할 수 없는 경우임

〈표 Ⅳ-3-2〉 분석대상(2차-3차 조사 결과 비교)

2차 조사 결과	3차조사 결과	인원(명)
조사 참여 (5,527)	참여(협조성향)	5,130
	미참여(비협조성향)	256
	미참여(사망, 장애등록취소, 시설입소 등(조사제외))	141
조사 미참여 (594)	참여(협조성향)	129
	미참여(비협조성향)	316
	미참여(사망, 장애등록취소, 시설입소 등(조사제외))	149
6,121	–	6,121

2차 조사 결과	3차 조사 결과 분석1 (wt3_1)	3차 조사 결과 분석2 (wt3_2)
➜ 분석대상	분석대상	분석대상
➜ 분석대상	분석대상	분석대상
분석대상		
분석대상		분석대상
분석대상		분석대상
분석대상 (26명)		
5,998	5,386	5,831

⇧
응답확률 추정시 모형에 포함

이때 특성 변수는 장애인삶 패널조사의 설계변수, 공표단위, 1차 년도 가중치 조정 변수 등을 종합적으로 고려한다. 표본설계시 적용된 변수를 기초로 실시하고, 무응답 조정시 추가적으로 고려할 변수에 대해 추가적 탐색을 수행한다. 분석을 통해 조사참여 여부와 패널에 대한 주요 특성들간의 연관성을 파악하고, 조사참여(최종 응답여부)에 영향을 미치는 특성 변수를 찾고자 한다. 분석은 무응답 표본의 가구 특성과 응답여부에 따른 패널 개인 및 가구, 금융변수와의 연관성을 파악하는 것을 중심으로 한다. 이를 위해 응답표본과 무응답표본에 대해 개인·가구 특성 변수로 패널의 성별, 연령대, 혼인상태, 교육수준, 가구주 여부 등의 변수와 금융변수인 자산, 부채, 소득분포를 비교한다. 응답여부에 따른 특성변수와의 연관성 분석은 범주형 변수의 경우, 카이제곱 검정(독립성)을 실시하고, 연속형 변수의 경우 t검정을 수행한다. 그리고 CHAID[50] 분석을 실시하여, 응답여부에 영향을 미치는 주요 특성 변수를 파악한다.

2) 패널 개인 및 가구특성

무응답 패널 개인 및 가구특성은 2차 조사 참여자 5,527명 중 3차 조사 참여 결과를 기반으로 분석제외 대상자를 제외한 5,386명에 대해 분석한다. 분석변수로는 패널 개인 특성으로 거주지역, 성별, 연령대, 혼인상태, 교육수준, 가구주 여부, 장애유형, 그리고 장애정도를 활용하였다. 가구특성으로는 가구원 규모와 주택유형, 그리고 주택점유형태를 활용하였다. 범주형 변수는 카이제곱 검정($\chi^2 - test$)을 실시하였으며, 연속형은 독립 t-test를 실시한다.

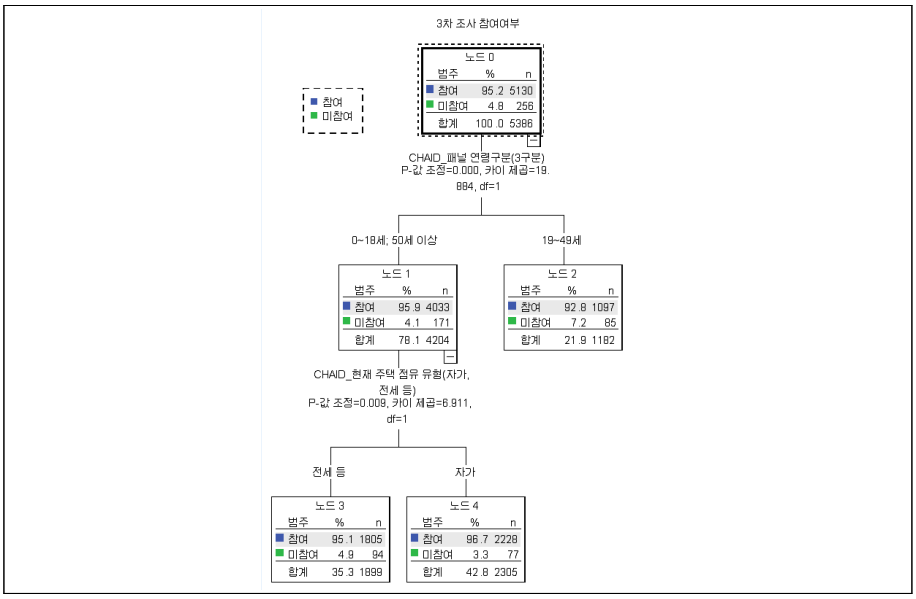
분석결과를 살펴보면, 개인특성으로는 거주지역, 연령대, 혼인상태, 교육수준, 장애유형이 3차년도 조사 참여에 대해 통계적으로 유의미한 차이가 있는 것으로 나타났다. 가구 특성에 대한 분석결과를 살펴보면, 가구원 수, 주택유형, 주택점유형태에서 통계적으로 유의미한 차이를 확인할 수 있다.

50) 이때 하위 노드 최소 표본 수와 트리 수를 각각 30으로 설정하였다.

〈표 Ⅳ-3-3〉 패널의 개인 및 가구 특성 분석 결과

패널 개인 및 가구특성 (분석대상: 5,386명)		df	χ^2	p-value
개인 특성	거주지역(17개 시도)	16	40.557	0.001***
	성별	1	0.006	0.939
	연령대	2	19.885	0.000***
	혼인상태(19세 미만, 미혼, 배우자 있음, 이혼, 별거, 사별)	5	19.768	0.001***
	교육수준(중졸미만, 중졸, 고졸, 전문대이상)	3	7.425	0.060*
	가구주 여부	1	1.159	0.282
	장애유형	14	41.067	0.000***
	장애정도	1	0.324	0.569
가구 특성	가구원 수(1인, 2인, 3인, 4인 이상)	3	8.179	0.042**
	주택유형(아파트, 연립 · 다세대, 단독 · 전원주택, 기타)	3	11.185	0.011**
	주택점유형태(자가, 임차 등)	1	9.293	0.002***

주 1) 분석대상: 5,386명
2) * : p<0.1, ** : p<0.05, *** : p<0.01



(투입변수) 거주지역, 성별, 연령대, 혼인상태, 교육수준, 가구주 여부, 장애유형, 장애정도, 가구원 수, 주택유형, 주택점유형태

[그림 Ⅳ-3-1] 개인 및 가구 특성의 CHAID 분석결과

3) 금융특성

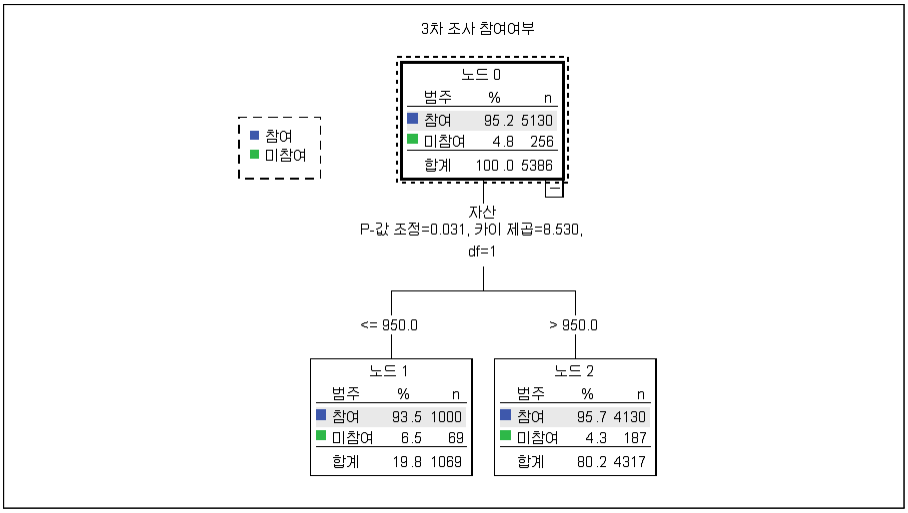
본 절에서는 금융특성과 응답여부에 따른 관계를 파악하기 위해 조사참여자와 조사 미참여자의 금융특성에 차이가 있는지를 분석하였다. 금융특성 변수는 근로소득, 사업/부업소득, 개인재산/금융/연금소득, 사적이전소득, 공적이전소득, 경상소득(계), 비경상소득, 월평균 가구 소득(계), 월 평균 가구 생활비(계), 자산, 부채를 활용하였다. 조사참여 여부에 따른 금융특성 변수의 차이를 분석한 결과 〈표Ⅳ-3-4〉와 같다. t-검정 결과로는 조사참여와 조사미참여가 금융특성 변수간의 유의미한 통계적 차이가 나타나지 않았다.

〈표 Ⅳ-3-4〉 패널의 금융특성 분석

항목	금융항목		평균(단위: 만원)		t	p-value
			조삼참여	조사미참여		
(월)소득	경상소득	근로소득	190.42	169.31	1.027	0.305
		사업/부업소득	37.11	52.94	-1.061	0.290
		개인재산/금융/연금소득	17.16	43.38	-.742	0.459
		사적이전소득	6.53	7.97	-.741	0.459
		공적이전소득	39.79	31.88	1.309	0.191
		경상소득(계)	291.02	305.48	-0.527	0.598
	비경상소득		1.77	3.96	-.738	0.461
	월평균 가구 소득(계)		292.79	309.44	-.607	0.544
월평균 가구 생활비(계)			183.37	191.79	-.868	0.386
자산			16297.15	14777.75	1.023	0.306
부채			3054.17	3747.69	-1.281	0.201

주: 1) 분석대상: 5,386명
2) 사업.부업소득, 개인재산.금융.연금소득, 비경상소득, 월 평균 가구생활비는 Levene test 결과 등분산 가정을 만족하지 않음(유의수준 5%)

CHAID 분석 결과를 살펴보면, 금융특성 변수에서 3차 조사 참여여부에 영향을 미치는 주요 변수는 자산인 것으로 나타났다.



(투입변수) 근로소득, 사업/부업 소득, 개인재산/금융/연금소득, 사적이전소득, 공적이전소득, 경상소득, 비경상소득, 월평균 가구소득, 월평균 가구생활비, 자산, 부채

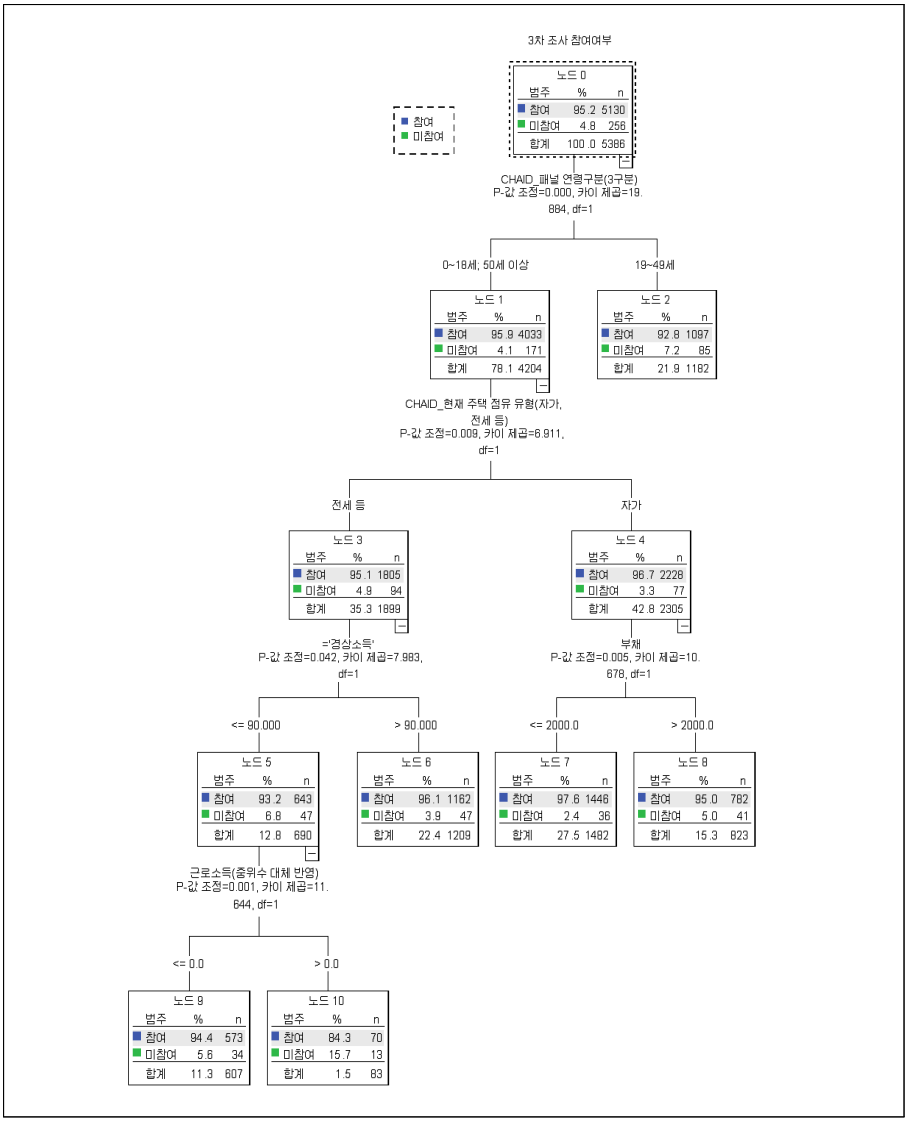
[그림 Ⅳ-3-2] 금융특성의 CHAID 분석결과

4) 영향요인 분석

CHAID 모형에 투입된 변수는 다음과 같다. 투입변수는 표본설계 변수를 기본으로 무응답 발생과 관련이 있을 것으로 판단되는 변수(지역, 주택유형, 가구주 정보, 금융 정보)를 추가로 탐색할 수 있도록 분석 모형에 투입하였다.

CHAID 모형 투입변수	• (패널 및 가구특성) 거주지역, 성별, 연령대, 혼인상태, 교육수준, 가구주 여부, 장애유형, 장애정도, 가구원 수, 주택유형, 주택점유형태 • (금융특성) 금융특성(근로소득, 사업/부업 소득, 개인재산/금융/연금소득, 사적이전소득, 공적이전소득, 경상소득, 비경상소득, 월평균 가구소득, 월평균 가구생활비, 자산, 부채)
분석 결과 모형 구성 변수	• 패널의 연령대, 교육수준, 가구원 수, 장애정도, 장애유형, 주택유형, 주택점유형태

3차년도 응답에 영향을 미치는 요인으로는 '주택점유형태', '가구월평균 소득'인 것으로 나타났다.



(투입변수) 거주지역, 성별, 연령대, 혼인상태, 교육수준, 가구주 여부, 장애유형, 장애정도, 가구원 수, 주택유형, 주택점유형태, 금융특성(근로소득, 사업/부업 소득, 개인재산/금융/연금소득, 사적이전소득, 공적이전소득, 경상소득, 비경상소득, 월평균 가구소득, 월평균 가구생활비, 자산, 부채)

[그림 Ⅳ-3-3] CHAID 분석결과

4 3차년도 가중치 처리 방안 비교 · 검토

3차년도 가중치 부여 대상은 3차년도 조사에 참여하여 응답완료한 모든 패널이다. 다음 〈표 Ⅳ-4-1〉에서 응답패턴1과 응답패턴3을 가지는 패널이다. 3차년도 응답결과에 따라 연속참여패널과 비연속패널 분류할 수 있는데, 연속참여 패널은 3차 조사를 기준으로 1차-3차 조사에 모두 참여한 패널로 응답패턴1을 의미한다. 응답패턴3은 3차년도 조사에 재진입한 패널로 직전년도 가중치가 존재하지 않는다. 따라서 3차 조사부터는 응답패턴1과 응답패턴3에 가중치를 부여하며, 해당차수 기준 연속응답패널과 해당차수 응답패널에 종단가중치가 부여된다. 결론적으로 두 종류의 가중치가 산출된다. 이를 통해 연속참여자를 대상으로 시계열 분석을 실시하고, 당해차수 조사참여자 현황을 파악할 수 있는 기초통계용 가중치로 활용할 수 있다.

본 절은 이렇게 응답패턴이 다른 경우 이들에 대한 가중치를 어떤 방법으로 부여할 것인가에 대해 살펴본다. 구체적으로는, 첫째, 기본적으로 가중치는 직전년도 가중치를 기반으로 산출되는데, 재진입 패널을 어떻게 처리할 것인가? 둘째, 연속참여 패널의 가중치 산출은 어떤 방법을 적용할 것인가? 에 대해 살펴보고자 한다.

〈표 Ⅳ-4-1〉 가중치 부여 대상

구분	계(명)	1차(2018)	2차(2019)	3차(2020)		
				응답여부	가중치 부여 대상	비고
응답패턴1	5,130	1	1	1	O	2차 가중치 없음
응답패턴2	397	1	1	0	–	
응답패턴3	129	1	0	1	O	
응답패턴4	465	1	0	0	–	
계(명)	6,121	6,121	5,527	5,259		

1) 산출하는 종단가중치의 종류

장애인삶 패널조사의 표본은 연차별로 조사 참여 및 미참여로 구분할 수 있다. 이때 3차년도 조사의 미참여 표본은 3차 조사에서 최초로 이탈한 패널과 2차 조사에서부터 이탈한 패널이 존재한다. 3차년도 조사의 참여 표본은 1차-3차 조사에 모두 참여한 연속참여패널과 2차년도 조사 미참여 후 3차년도 조사에 참여하는 재진입(re-entrance)패널(또는 비연속참여패널)이 존재한다. 따라서 장애인삶 패널조사는 이러한 조사 참여 유형에 따른 두 종류의 가중값을 산출하며, 부여 대상은 〈표 Ⅳ-4-2〉와 같다. 하나는 1차-3차 조사에 연속응답한 대상자에게 부여하는 가중치(wt3_1)이며, 다른 하나는 해당차수 응답한 전체 대상자에게 부여하는 가중치(wt3_2)이다.[51] 연속참여패널에 부여된 전자의 가중치(wt3_1)는 시계열분석에 적합한 3차년도 종단가중치이다. 여기에 재진입 패널은 포함되지 않는다. 후자는 3차년도 조사에 참여한 모든 패널, 즉 재진입 패널을 포함하여 산출한 종단가중치로 기초통계분석을 위한 것이다.

장애인삶 패널조사는 2015년~2017년 사이에 장애등록을 한 장애인을 목표 모집단으로 정의하고, 이들을 매년 추적조사하고 있다. 따라서 본 조사는 시간 경과에 따른 모집단내 변화를 설명하는데 의미가 있다. 다시 말해 초기 구축한 표본이 당해 연도 기점에서 초기 등록장애인을 대표한다고 보는 것은 무리가 있다. 따라서 모집단의 변화와 모집단내 변화가 동일한 의미를 가지며,[52] 본 패널조사에서 wt3_2는 엄밀한 의미에서의 횡단가중치로 정의하기는 어렵다. 이러한 이유로 장애인삶 패널조사는 시계열분석에 적합한 가중치를 연속참여패널에 부여(wt3_1)하고, 3차년도 기준 참여한 모든 패널에 가중치를(wt3_2)를 부여한다. 이때 〈표 Ⅳ-4-2〉의 「가중치유형 ②: wt3_2」에서 모집단의 변화를 의미하는 기준 모집단은 패널의 구축 모집단(1차 조사)인 2015-2017년 장애등록을 한 등록장애인 모집단과 동일하게 된다.[53]

51) 이정화, 신슬비. (2017). 제4차 산재보험패널조사 기초분석 보고서(2017). 근로복지공단
52) 김현지, 김태용, 이혜수. (2020). 장애인삶 패널조사를 위한 종단가중치 작성 방안 연구. 한국장애인개발원.
53) 김현지, 김태용, 이혜수. (2020). 장애인삶 패널조사를 위한 종단가중치 작성 방안 연구. 한국장애인개발원.

〈표 Ⅳ-4-2〉 응답패턴별 산출하는 종단가중치의 종류

기준모집단: 패널구축시 모집단 N=251,277

가중치 유형 ①: wt3_1 (연속응답자)

응답패턴	1차(2018)	2차(2019)	3차(2020)
패턴1	1	1	1 → 종단
패턴2	1	1	0
패턴3	1	0	1
패턴4	1	0	0

종단가중치 부여	인원(명)	5,527	5,130 (wt3_1)
기준모집단	계(명)	251,277	251,277

가중치 유형 ②: wt3_2 (해당차수 전체 응답자)

응답패턴	1차(2018)	2차(2019)	3차(2020)
패턴1	1	1	1
패턴2	1	1	0
패턴3	1	0	1
패턴4	1	0	0 ↓ (기초 통계용)

종단가중치 부여	인원(명)	5,527	5,259 (wt3_2)
기준모집단	계(명)	251,277	251,277

〈표 Ⅳ-4-3〉 3차조사 결과에 따른 산출하는 종단가중치의 종류 및 부여 대상

2차 조사 결과	3차조사 결과	인원(명)		2차 조사 가중치	3차 종단가중치 wt3_1	3차 종단가중치 wt3_2
조사 참여 (5,527)	참여(협조성향)	5,130	➜	가중치	가중치	가중치
	미참여(비협조성향)	256	➜	가중치		
	미참여(사망, 장애등록취소, 시설입소 등(조사제외))	141		가중치		
조사 미참여 (594)	참여(협조성향)	129				가중치
	미참여(비협조성향)	316				
	미참여(사망, 장애등록취소, 시설입소 등(조사제외))	149				
6,121	–	6,121		5,998	5,386	5,386

2) 종단가중치(3차) 처리를 위한 검토 방안

패널조사는 시간에 따라 추적조사 하는 특성으로 인해 조사차수가 증가하면서 다양한 사유로 인해 패널 이탈이 발생한다. 패널 이탈로 인해 발생한 무응답은 단위무응답으로 대표성 상실, 추정치의 편향 등을 야기하게 된다. 따라서 무응답을 반영하여 가중치를 산출하는데, 이때 이전 조사 자료에서 얻은 표본에 대한 일부 정보(김재광 등, 2004)[54]의 활용이 가능하다.

본 절에서는 무응답을 및 응답패턴을 고려하여 3차년도 종단가중치 처리를 위해 최근가중치 승계 방안과 응답확률 추정 모형을 검토한다. 3차년도 가중값 조정은 2차년도 가중치를 기본가중치로 고려하여 3차년도 종단가중치를 산정한다. 가중치 조정 대상은 먼저 2차년도 응답한 개인이 3차년도에 지속적으로 응답한 패널과 3차년도에 응답하지 않은 개인들을 토대로 응답확률을 추정하여 3차년도 응답개인들의 종단가중치로 조정한다. 3차년도 가중치 부여 대상은 3차년도 조사에 참여하여 응답완료한 모든 패널 이며(〈표 IV-4-1〉), 응답패턴에 따라 두 종류의 가중치가 산출된다. 패널조사의 종단가중치는 이전 차수의 가중치를 기반으로 하지만 2차 조사에 미참여 표본의 경우에는 이전차수인 2차 가중치가 존재하지 않는다. 향후 자료 사용 측면을 고려하여 장애유형은 15개 범주로 적용하고, 레이킹 비 조정시에도 15개 범주를 적용 하였다. 최종 작성되는 가중치의 최소값은 1이 되도록 하였다. 여기서 무응답 표본은 2장에서 논의한 정의에 따라 1차 조사에 참여하여 패널로 구축되었고, 2차 조사에서 조사 가능하였으나 최종적으로 조사에 참여하지 않은 표본이다. 검토한 방안은 〈표 Ⅳ-4-6〉과 같다.

응답확률 추정 모형에 투입될 설명변수는 분석상의 일관성을 확보하고자 동일한 설명변수를 적용하는 방법이 활용되기도 한다. 그러나 개념상 2차 조사의 응답지속여부와 3차 조사에서의 응답지속여부의 결정요인은 상이할 수 있기 때문에 조사시점별 응답모형을 적용하는 것이 타당할 수 있다(주재선 외, 2012).[55] 응답확률 추정 시의 주요 이슈는 응답확률 추정을 위한 통계적 모형과 추정모형에 투입(input)할 설명변수가 무엇이냐 하는 것이다(김현지 외, 2020에서 재인용). 박민규 외 (2011)는 종단가중치 산출을 위해 주요 설명변수별 3차년도 응답자와 무응답자간 차이를 비교·분석하였다. 그리고 변수선택 방법으로 로지스틱 회귀모형을 적용(stepwise)

54) 김재광, 한근식, 윤연옥. (2004). 가계조사 무응답 처리기법 연구. 통계연구. 9(1), 79-102.
55) 주재선 외. (2012). 여성가족패널조사. 한국여성정책연구원.

하여 설명변수의 동시 예측력을 비교·분석하여 모든 차수에서 응답에 영향을 미치는 주요 변수를 선택하였다. 이에 장애이삶 패널조사는 가중치 작성의 일관성, 조사결과의 공표 등을 고려하여 1차 년도와 마찬가지로 장애특성 및 인구특성 변수는 일관되게(장애유형, 성별, 장애정도, 연령대)유지하고자 한다. 그리고 장애유형, 성별, 장애정도, 연령대 변수를 조정 변수로 레이킹 비 조정하였으며, 이때의 기준 모집단은 1차 년도 패널 구축 당시 모집단을 기준으로 한다.[56)]

가중치 산출을 위해 어떤 방법을 적용할 것인가? 본 조사에서는 최근가중치 승계와 응답확률 추정 모형 적용을 검토한다. 응답확률 추정 모형 적용시 투입할 설명변수를 선택하고, 이를 위해 검토한 방안은 다음 3가지이다. 이를 통해 최종적으로 적용할 방법을 마련하고자 한다. 첫 번째 방안은 가장 최근년도 부여된 가중치를 기본가중치로 승계하여 3차년도 무응답 조정 후 레이킹비 조정으로 가중치를 산출하는 것이다. 방안3은 3차 응답자를 기준으로 3차 응답여부에 대한 응답확률을 추정한다. 방안2는 최근가중값 승계와 응답확률 추정 모형을 적용하는 방안을 혼합한다. 이는 연속참여 패널에 대해서는 직전가중치를 기반으로 3차 응답여부에 대한 응답확률모형을 추정하고, 재진입 패널은 가장 최근가중값을 승계하여 최종 가중치를 산출하는 것이다. 최근가중치 승계, 응답확률 추정 모형으로 본 조사에서 검토한 방안을 구체적으로 살펴보면 다음과 같다.

◦ (방안1) 가장 최근가중치 승계

방안1은 가장 최근가중치를 승계 받아 2차 조사에 미참여한 경우 1차 가중치를 기본가중치로 가진다.

56) 이에 대해서는 「김현지, 김태용, 이혜수. (2020). 장애인삶 패널조사를 위한 종단가중치 작성 방안 연구. 한국장애인개발원」 3장 소결 및 시사점 참고한다.

〈표 Ⅳ-4-4〉 (방안1) 가장 최근가중치 승계

가중치				
1차	2차	3차	3차 조사참여 여부	비고
●	◎ →	◎	참여	가장 최근가중치 승계
●	◎	X	미참여	–
●	X →	●	참여	가장 최근가중치 승계
●	X	X	미참여	–

◦ (방안2) 2차년도 응답자의 3차년도 응답여부 기준으로 응답확률 추정

방안2는 2차 조사에 참여한 경우 응답확률 추정 결과를 토대로 가중치가 반영되며, 2차에 미참여한 경우에만 1차 가중치를 기본가중치로 가진다.

〈표 Ⅳ-4-5〉 (방안2) 2차년도 응답자의 3차년도 응답여부 기준으로 응답확률 추정

가중치				
1차	2차	3차	3차 조사참여 여부	비고
●	◎	○	참여	응답확률 추정(로지스틱 회귀모형)
●	◎	X	미참여	
●	X →	●	참여	가장 최근가중치 승계
●	X	X	미참여	–

주) 3차 응답기준으로 조사불가(사망, 장애등록취소)는 제외

방안2에서는 응답확률 추정시(로지스틱 회귀모형) 패널의 고유특성인 '장애유형', '장애정도', '성별', '연령대'와 3차년도 응답에 영향을 미치는 '주택점유형태', '가구월평균 소득'을 활용한다. 2차년도 조사의 경우 '주택유형', '가구규모', '월평균 생활비'가 응답에 영향을 미치는 변수였다. 시간에 따라 응답에 영향을 주는 변수는 상이 할 수 있다는 가정 하에서 검토한 방안으로 가장 최신 시점의 정보(2차년도 조사결과)를 활용하여 응답확률을 추정한다.

방안3은 방안2와 동일한 형태이지만 3차 응답자를 기준으로 3차 응답여부에 대한 응답확률을 추정하는 방안이다. 이때 무응답 표본은 가중치 작성 방법의 일관성 등을 고려하여, 2차년도에 정의한 무응답 기준에 따라 조사제외 대상은 무응답 패널으로 정의하지 않았다. 조사가능 패널 중 조사미참여를 무응답으로 정의하여, 최종적으로 가중치 부여 대상은 연속참여패널 5,130명(wt3_1)과 3차년도 조사참여 패널 5,259명(wt3_2)이다. 최종 작성되는 가중치의 최소값은 1이 되도록 하였으며, 최종적으로 검토한 방안은 〈표 Ⅳ-4-6〉과 같다. 레이킹 비 조정시 기준 모집단은 초기 구축한 표본의 모집단인 2015-2017년 등록한 초기등록장애인을 모집단으로 한다.

〈표 Ⅳ-4-6〉 종단가중치(3차) 산출 방안 검토

방안		종단가중치 유형 ①: wt3_1 (연속응답자: 5,130명)	종단가중치 유형 ②: wt3_2 (기초통계용: 5,259명)
최근 가중치 승계	방안 1	• 가장 최근 가중치를 원가중치로 무응답 보정 후 레이킹 비 조정 → 레이킹 비 조정(조정변수: 지역, 장애유형, 장애정도, 성별, 연령)	• 5,130명: 가장 최근 가중치(전년도 가중치)를 원가중치 • 129명(재진입): 가장 최근 가중값 승계 → 레이킹 비 조정(조정변수: 지역, 장애유형, 장애정도, 성별, 연령)
응답 확률 모형 + 최근 가중치 승계	방안 2	• 2차 응답자 기준으로 3차 응답여부에 대한 응답확률모형 적용(2차와 동일) → 레이킹 비 조정 – 응답확률모형: (설명변수) 지역, 장애유형, 장애정도, 성별, 연령, 주택점유형태, 가구월평균소득 – 레이킹 비 변수: 지역, 장애유형, 장애정도, 성별, 연령	• 5,130명: 2차 응답자 기준으로 3차 응답여부에 대한 응답확률모형 적용(2차와 동일) • 129명(재진입): 가장 최근 가중값 승계 → 레이킹 비 조정(조정변수: 지역, 장애유형, 장애정도, 성별, 연령)
응답 확률 모형	방안 3	• 3차 응답자 기준으로 3차 응답여부에 대한 응답확률모형 적용(2차와 동일) → 레이킹 비 조정 – 응답확률 추정 모형: (설명변수) 장애유형, 장애정도, 성별, 연령, 주택형태, 가구월평균생활비, 가구원수 – 1차 자료(모든 개체에 대해 값이 존재)	• 종속변수: 3차 응답여부 • 설명변수: 지역, 장애유형, 장애정도, 성별, 연령, 주택형태, 가구월평균생활비, 가구원 수 – 1차 자료(모든 개체에 대해 값이 존재)

3) 검토 방안별 비교 · 검토

본 절에서는 장애인삶패널조사 종단가중치 부여를 위해 앞서 살펴본 검토방안별 가중치 분포를 비교 · 검토하고자 한다. 먼저 로지스틱 회귀분석을 통해 응답확률[57]을 추정하여 종단가중치를 산출하는 방안2와 방안3을 살펴본다. 방안2에서는 가장 최근 년도 데이터를 활용하여 로지스틱 회귀모형을 적용한다. 설명변수는 앞 절의 응답자와 무응답자간 특성 분석 결과를 바탕으로 활용하였다. 이때 종속변수는 조사참여 여부로 구분되는 지시변수이다. 3차년도 응답할 확률에 대한 로지스틱 회귀모형이 적합하고, 이를 통해 산출되는 응답확률(propensity score)을 가중치 보정에 적용한다. 이때 조사 미참여 사유가 사망(자연소멸)으로 향후 조사부터는 패널로 참여할 수 없는 경우와 장애등록취소, 시설(병원)입소로 조사불가로 미참여한 패널은 종단가중치 작성을 위한 응답확률 추정 모형에서 제외하였다.

$$\circ\ 3\text{차년도의 응답여부} = \begin{cases} 1, & \text{응답} \\ 0, & \text{무응답} \end{cases}$$

〈표 Ⅳ-4-7〉 방안2의 설명변수

변수	변수유형	범주	내용
장애유형	범주형	15	◦ 지체, 시각, 청각, 언어 ,지적, 자폐성, 뇌병변, 정신 신장, 심장, 호흡기, 간, 안면, 장루·요루, 뇌전증
장애정도	범주형	2	◦ 중증, 경증
성별	범주형	2	◦ 남자, 여자
연령대	범주형	3	◦ 만18세 이하, 만19-49세, 만50세 이상
주택점유형태	범주형	2	◦ 자가, 전세 등
가구월평균소득	범주형	4	◦ 100만 미만, 100만 이상-200미만, 200만 이상-300미만, 300만 이상

57) 응답경향점수(propensity score)라고도 불린다.

〈표 Ⅳ-4-8〉 방안2의 응답확률 추정 모형(로지스틱 회귀)

변수	범주	추정치	표준오차	오즈비	p-값
Intercept	–	3.1684	0.2616	–	<.0001
연령대	만0세–만18세	–0.1949	0.2171	0.823	0.3694
(기준: 만50세 이상)	만19세–만49세	–0.6224***	0.1538	0.537	<.0001
장애유형	간장애	–0.4871	0.3579	0.614	0.1736
(기준: 지체장애)	뇌병변장애	–0.1489	0.2358	0.862	0.5277
	뇌전증장애	0.3753	0.5424	1.455	0.4890
	시각장애	–0.3157	0.2370	0.729	0.1828
	신장장애	–0.2914	0.2685	0.747	0.2779
	심장장애	0.5149	0.6131	1.673	0.4010
	안면장애	–2.0631***	0.4808	0.127	<.0001
	언어장애	–0.1450	0.3149	0.865	0.6451
	자폐성장애	–0.0793	0.4520	0.924	0.8608
	장루.요루장애	0.4058	0.6107	1.500	0.5064
	정신장애	0.0664	0.3378	1.069	0.8443
	지적장애	–0.0118	0.3188	0.988	0.9705
	청각장애	0.1532	0.2698	1.166	0.5702
	호흡기장애	0.9987	0.7382	2.715	0.1761
장애정도 (기준: 중증)	경증	0.00710	0.1514	1.007	0.9626
성별 (기준: 남자)	여자	–0.00355	0.1306	0.996	0.9783
주택점유형태 (기준: 자가)	전세 등	–0.2748*	0.1428	0.760	0.0544
가구월평균소득 (기준: 100만 미만)	100만 이상 –200만 미만	0.4418**	0.1977	1.555	0.0254
	200만 이상 –300만 미만	0.3092	0.2032	1.362	0.1280
	300만 이상	0.3571*	0.1937	1.429	0.0652

* p>0.1 ** p<0.05 *** p<0.01. 설명변수: 2차 조사 자료.

방안3에서는 3차 조사가능 패널의 3차 응답여부를 기준으로 응답확률을 추정하였으며, 로지스틱 회귀모형 적용을 위해 1차년도 데이터를 활용하였다. 5,386명이다. 응답확률 추정시 1차 조사 자료를 사용하였으며, 2차 조사의 가중치 산출과 동일한 방법을 적용한다.

〈표 Ⅳ-4-9〉 방안3의 설명변수

변수	변수유형	범주	내용
장애유형	범주형	15	◦ 지체, 시각, 청각, 언어 ,지적, 자폐성, 뇌병변, 정신, 신장, 심장, 호흡기, 간, 안면, 장루·요루, 뇌전증
장애정도	범주형	2	◦ 중증, 경증
성별	범주형	2	◦ 남자, 여자
연령대	범주형	3	◦ 만18세 이하, 만19-49세, 만50세 이상
주택유형	범주형	4	◦ 아파트, 단독 · 전원주택, 연립 · 다세대, 기타
가구규모	범주형	4	◦ 1인가구, 2인가구, 3인가구, 4인 가구 이상
가구월평균 생활비	범주형	4	◦ 100만 미만, 100만 이상-200미만, 200만 이상-300미만, 300만 이상

$$\circ\ 3\text{차년도의 응답여부} = \begin{cases} 1, & \text{응답} \\ 0, & \text{무응답} \end{cases}$$

〈표 Ⅳ-4-10〉 방안3의 응답확률 모형(로지스틱회귀모형)

변수	범주	추정치	표준오차	오즈비	p-값
Intercept	-	2.601	0.348	-	<.0001
연령대 (기준: 만50세 이상)	만0세-만18세	-0.153	0.232	0.859	0.511
	만19세-만49세	-0.617***	0.155	0.54	<.0001
장애유형 (기준: 지체장애)	간장애	-0.534	0.358	0.586	0.137
	뇌병변장애	-0.177	0.237	0.838	0.455
	뇌전증장애	0.231	0.544	1.26	0.671
	시각장애	-0.346	0.237	0.708	0.145
	신장장애	-0.251	0.269	0.778	0.350

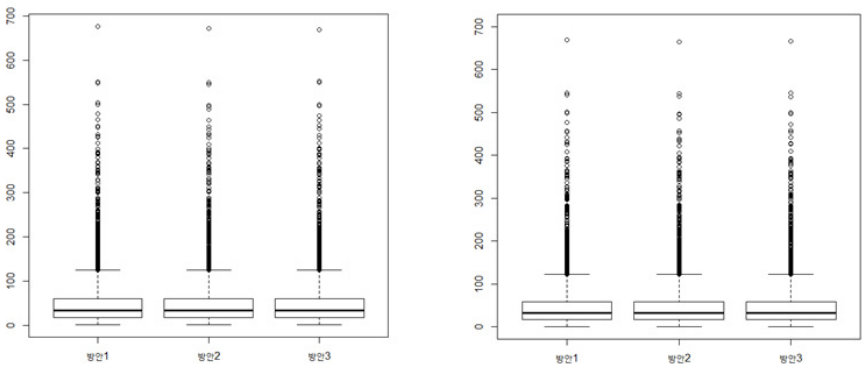
변수	범주	추정치	표준오차	오즈비	p –값
	심장장애	0.473	0.614	1.605	0.441
	안면장애	–2.114***	0.484	0.121	<.0001
	언어장애	–0.194	0.316	0.824	0.540
	자폐성장애	–0.091	0.454	0.913	0.842
	장루.요루장애	0.377	0.611	1.458	0.538
	정신장애	–0.066	0.338	0.936	0.846
	지적장애	–0.078	0.320	0.925	0.808
	청각장애	0.131	0.270	1.14	0.628
	호흡기장애	0.964	0.738	2.621	0.192
장애정도 (기준: 중증)	경증	0.049	0.152	1.05	0.749
성별 (기준: 남자)	여자	–0.018	0.131	0.982	0.891
주택점유형태 (기준: 자가)	단독.전원주택	0.568*	0.323	1.765	0.078
	아파트	0.381	0.294	1.463	0.196
	연립.다세대	0.325	0.303	1.384	0.284
가구규모 (기준: 1인 가구)	2인가구	0.454**	0.201	1.574	0.024
	3인가구	0.705***	0.245	2.024	0.004
	4인가구	0.534**	0.253	1.706	0.035
가구월평균생활비 (기준: 100만 미만)	100만 이상–200만 미만	–0.062	0.190	0.94	0.744
	200만 이상–300만 미만	–0.118	0.246	0.888	0.631
	300만 이상	–0.360	0.265	0.698	0.175

* p>0.1 ** p<0.05 *** p<0.01. 설명변수: 1차조사 자료.

3차년도 종단가중치는 무응답을 보정한 기본가중치 산출 후 레이킹 보정 과정을 통해 작성하였다. 기본가중치는 조사대상자의 개인 특성 및 금융 특성 등을 기반으로 응답확률 추정 모형을 적용하여 응답확률을 추정한 후, 그 역수를 무응답 보정하였다. 이렇게 산출한 가중치합계는 모집단의 총계와 일치하지 않을 수 있으며, 이는 레이킹 비 보정하여 최종 종단가중치를 산출하였다.

〈표 Ⅳ-4-11〉 검토방안별 종단가중치의 $\sqrt{1+cv^2}$ 및 가중치 분포

	1차 조사 가중치	wt3_1 (연속응답자)			wt3_2 (3차기준 응답자)		
		방안1	방안2	방안3	방안1	방안2	방안3
		최근 가중치 승계	로지스틱 회귀모형	로지스틱 회귀모형	최근가중치 승계	로지스틱 회귀모형 (재진입: 가장 최근가중치 승계)	로지스틱 회귀모형 (3차년도 기준 분석대상 패널로 응답확률 모형: 5,831명)
가중치 인원	6,121	wt3_1	wt3_1	wt3_1	wt3_2	wt3_2	wt3_2
		5,130	5,130	5,130	5,259	5,259	5,259
평균	41.052	48.982	48.982	48.98187	47.780	47.780	47.780
표준편차	47.966	54.424	54.402	54.366	53.449	53.440	53.426
$\sqrt{1+cv^2}$	1.538	1.4948	1.4945	1.4940	1.5004	1.5003	1.5001
가중치합계	251,277	251,277	251,277	251,277	251,277	251,277	251,277



[그림 Ⅳ-4-1] 검토방안별 가중치 분포

위 세 가지 방안에 대해 가중치 분포를 살펴보면, 검토방안별로는 통계적으로 유의한 차이가 발견되지는 않는다. Unequal weighting effect값도 유의한 차이를 보이지 않고 있다. 이에 검토 방안 중 어떤 방안을 적용하여도 큰 무리가 없을 것이라고 판단된다. 다음절에서 층화변수별 가중치 분포, 주요문항별 상대표준오차를 검토하고 최종적으로 적용 방안을 선정 하고자 한다. 방안2는 최신시점의 자료를 활용하면서 가중치 산출 방법의 일관성을 확보할 수 있다는 점에서 최종적으로 응답확률 추정 모형과 최근가중값을 승계하는 방법을 혼합하는 방안2를 적용하는데 큰 무리가 없을 것으로 판단된다.

〈표 Ⅳ-4-12〉 3차년도 종단가중치(wt3_1) 기술통계량_방안2

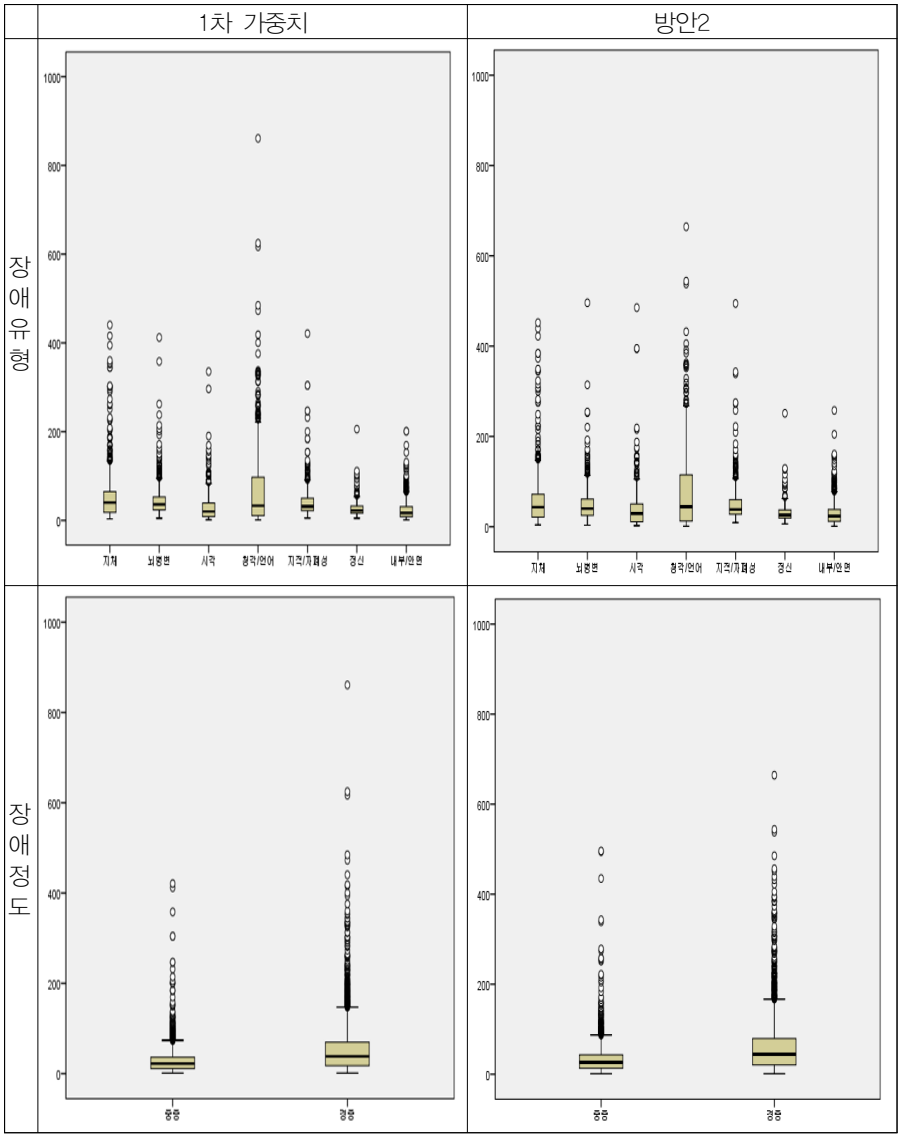
관측수	5,130	가중합	5,130
평균	48.9818713	관측값 합	251277
표준 편차	54.4020015	분산	2959.57777
왜도	3.47891511	첨도	18.7690538
제곱합	27487692.1	수정 제곱합	15179674.4
변동계수	111.065584	평균의 표준 오차	0.7595497
중위수	33.50423	표준 편차	54.40200
사분위수 범위	43.30916	범위	671.12120

주) 연속참여패널: 1~3차

〈표 Ⅳ-4-13〉 3차년도 종단가중치(wt3_2) 기술통계량_방안2

관측수	5,259	가중합	5,259
평균	47.7803765	관측값 합	251277
표준 편차	53.4403244	분산	2855.86827
왜도	3.52119892	첨도	19.2378638
제곱합	27022265	수정 제곱합	15016155.4
변동계수	111.845758	평균의 표준 오차	0.73691518
중위수	32.56844	표준 편차	53.44032
사분위수 범위	42.21996	범위	663.30051

주) 3차년도 조사참여자



[그림 Ⅳ-4-2] 3차년도 종단가중치 분포 wt3_2 (5,259명)

5 주요항목 분석

본 절에서는 표본조사의 통계 품질을 가늠하는 측도로 주로 활용되는 추정량의 상대표준오차를 분석한다. 주요항목은 1차년도 가중치 산출시 선정한 주요항목을 기준으로 상대표준오차를 산출하고 검토방안별 비교·분석한다. 본 절에서 검토한 주요문항은 이전 차수와 동일한 월평균 가구소득, 노후생활에 필요한 1달 생활비, 장애수용, 일상생활존중정도 이다. 이때 일상생활존중정도는 비율추정 결과이다.

〈표 Ⅳ-5-1〉 변수유형별 주요 문항

변수유형	문항	조사영역(세부영역)
연속형 변수	• 월평균 가구소득	장애수용 및 변화(장애수용, 심리상태, 사회관계, 경제상황, 주거)
	• 노후생활에 필요한 1달 생활비	사회참여(장애관련 서비스, 인권, 결혼생활 및 성, 노후, 삶의 만족도)
척도 변수(평균) : 세부문항으로 구성	• 장애수용	장애수용 및 변화(장애수용, 심리상태, 사회관계, 경제상황, 주거)
단일 범주형 변수(비율)	• 일상생활존중정도	사회참여(장애관련 서비스, 인권, 결혼생활 및 성, 노후, 삶의 만족도)

주) 김현지, 김태용. (2019). 장애인삶 패널조사 가중치 산정 방안 연구. 에서 발췌

앞서 검토한 방안별 종단가중치 작성 후, 주요 문항들에 대한 상대표준오차를 산출한 결과, 전국 단위 상대표준오차는 5% 미만으로 나타났다. 상대표준오차는 전반적으로 만족할 만한 수준이며, 장애유형, 장애정도별 상대표준오차 역시 20%를 넘지 않았다. 또한 검토방안별 모집단 전체 및 관심 도메인 추정량의 상대표준오차는 유의한 차이를 보이지 않는 것으로 보아, 전반적으로 만족할 만한 수준으로 보인다.

따라서 분석결과를 종합해보면, 방안2를 적용하는 것이 무리가 없을 것으로 판단된다. 장애인삶 패널조사의 추출단위 및 주요 분석 단위가 개인임을 고려할 때 이전 차수와의 분석상 일관성 확보 차원을 감안하여 응답확률 추정 모형으로 로지스틱 회귀모형을 적용하는 방안2를 적용하고자 한다.

〈표 Ⅳ-5-2〉 주요문항별 상대표준오차

구분	wt3_1			wt3_2		
	방안1	방안2	방안3	방안1	방안2	방안3
월평균 가구소득	1.6644	1.6686	1.6664	1.6487	1.6532	1.6484
노후생활에 필요한 1달 평균 생활비	0.9752	0.9747	0.9761	0.9700	0.9694	0.9706
장애수용(평균)*	0.4107	0.4103	0.4101	0.4068	0.4065	0.4065
일상생활존중받는 정도(존중받고 있는 비율)**	1.7016	1.7025	1.7009	1.6878	1.689	1.6867

* 주 1) 12개 세부 항목의 전체 평균을 사용하여 산출(일부 문항 역코딩)
2) 4점 리커트 척도: 전혀 그렇지 않다. 거의 그렇지 않은 편이다. 그런 편이다. 매우 그렇다.

** 주 1) 일상생활존중받는 정도: 전혀 존중받지 못한다. 거의 존중받지 못하는 편이다. 존중받고 있는 편이다. 매우 존중받고 있다.
2) 존중받고 있는 비율은 '존중받고 있는 편이다'와 '매우 존중받고 있다'를 합친 비율이다.

한국장애인개발원
www.koddi.or.kr

Ⅴ. 결 론

1 결론

패널조사는 종단연구 방법의 하나로 시간 경과에 따른 개인의 변화를 통계적으로 분석하여 모집단의 변화양상에 대한 이해 또는 모집단 변화의 원인에 대한 인과적 설명을 목적으로 필요성이 증대되고 있다(이경상 외, 2008)[58]. 따라서 패널의 추적 관리가 패널조사에서 중요한 요소 중 하나이다. 본 연구에서는 선행자료 분석, 3차 조사 참여자 특성 분석, 패널이탈 및 응답패턴분석을 실시하였다. 또한 응답여부에 영향을 미치는 영향 요인 분석, 그리고 최근가중치 승계와 응답확률 추정 모형을 기반으로 종단가중치(3차년도) 산출 방안을 검토하였다. 본 연구의 초점은 장애인삶 패널조사의 패널이탈 특성을 파악하여이탈 패널의 정보를 토대로 패널 관리의 기초자료를 확보와 종단가중치 조정에 대한 방안을 마련하는데 있다.

장애인삶 패널조사는 1차 조사와 패널 구축이 동시에 이루어져 총 6,121명의 장애인 패널을 구축하였다. 2차 조사(2019)는 5,527명 조사완료, 3차 조사(2020)는 5,259명 조사 완료하여 표본유지율은 85.92%이다. 3차년도 조사에서는 2차년도 조사에 참여하지 않았으나 3차년도 조사에 재진입한 패널이 나타난다. 이는 시간의 경과에 따라 동일 대상을 추적 조사하는 패널조사의 특성에 기인한 것으로 이탈과 진입을 반복하는 재진입 표본이다. 패널조사로부터 생성된 데이터 분석을 위해서는 이러한 특성을 반영한 가중치 산출이 필요하다. 이러한 가중치를 활용하여 유한모집단의 종단분석을 실시하는 것이 패널의 대표성과 추정 결과의 강건성 유지를 위해 바람직하다.[59]

종단가중치는 표본 이탈로 인해 발생하면서 나타나는 추정치의 편향에 대응하는(Uhirg, 2008: 이상호와 이혜정, 2008에서 재인용; 홍민기와 최효미, 2014; 하형석 외, 2018)[60] 방법이다. 하지만 이러한 통계적 접근 외에도 조사과정에서 양적·질적으로 패널 이탈의 최소화를 위한 노력이 병행되어야한다 (이상호, 2005). 이를 위해서는 패널 이탈 연구 등 데이터 품질 분석에 기반을 두어 이탈을 최소화하여 패널을 관리할 수 있는 다양한 접근이 필요하다. 이번 연구를 기점으로 데이터 품질 관리를 포함한 다양한 통계적 이슈를 다룰 수 있는 연구가 지속될 필요가 있다.

---

58) 이경상 외. (2008). 한국 청소년 패널조사(KYPS) VI. 한국청소년정책연구원.
59) 박민규, 김사라. (2015). 한국노동패널조사 자료의 분석을 위한 패널 가중치 산출 및 사용방안 사례 연구. 조사연구. 16(1), 49-71.
60) 하형석 외. (2018). 한국아동·청소년패널조사IX. 한국청소년정책연구원.

장애인삶 패널조사의 3차년도 종단가중치 산출을 위해 응답패턴이 다른 연속참여패널과 재진입 패널의 가중치 처리 방안을 검토하였다. 이는 동일 대상을 반복적으로 조사하는 특성에 따른 것으로 응답패턴에 따른 두 종류의 가중치를 산출하였다. 첫 번째 종류는 시간 경과에 따른 유한모집단의 동태적 분석에 필요한 가중치이다. 두 번째는 당해 연도 조사 자료를 이용한 통계 산출 및 분석을 위한 가중치이다. 검토방안으로는 ① 최근가중치 승계, ② 응답확률 추정 모형+최근가중치 승계, ③ 응답확률 추정 모형을 검토하였으며, 향후 조사 차수에도 일관되게 적용할 수 있는 방안을 도출하고자 하였다.

최종적으로 장애인삶 패널조사의 3차년도 종단가중치는 응답확률 추정 모형과 최근가중치 승계 방법을 함께 적용한 방안2를 적용한다. 최종가중치의 선택은 통계적 추정과 함께 공표단위, 데이터 활용도 등을 포함하여 결정되어야 하며, 응답확률 추정 모형과 최근가중치 승계 방법을 적용한 방안2 역시 이러한 관점에서 결정하였다. 5,130명의 연속참여 패널은 응답확률 추정 모형으로 응답확률 추정 후, 무응답 보정계수를 곱하여 가중치를 부여 하였으며, 재진입 패널은 최근가중치를 승계하였다. 그렇게 산출한 가중치의 합계는 모집단 총계와 다를 수 있으며, 이는 레이킹 비 조정하였다. 레이킹 비 조정은 2차년도와 동일한 방식을 적용하였으며, 레이킹 비 보정시 기준이 되는 모집단은 초기 구축시의 모집단을 적용한다. 방안2는 3차년도 조사참여 결과를 바탕으로 이전 조사 차수에 참여하지 않은 패널에 대해서는 가장 최근가중치를 승계할 수 있으며, 연속참여자에 대해서는 응답확률 추정 모형을 적용하여 무응답 보정 가중치를 산출함으로서 가중치 산출 방법의 일관성을 확보할 수 있다. 또한 이전 차수 조사에서 이탈 후 당해차수 조사에 재진입 하여 이전년도 가중치가 존재하지 않는 패널에 대해서는 설계 시점의 추출률에 기반한 최근가중치 승계를 적용함으로서 대표성을 최대한 반영한다는 특징이 있다. 재진입한 표본은 향후 차수에서도 발생 가능하다. 조사 차수별 응답 여부에 영향을 주는 변수들은 달라질 수 있을 것이다. 종단가중치가 표본 마모로 인한 모집단 대표성 문제를 완전히 해결해 줄 수 있는 것은 아니다(김기헌, 2003). 따라서 표본 이탈 연구 등 패널마모를 감소하기 위한 다각적인 노력이 지속적으로 필요하다.[61]

61) 김기헌. (2003). 한국노동패널조사(KLIPS)의 2003년 가중치.

한국장애인개발원
www.koddi.or.kr

참고문헌

참고문헌

강석훈. (2000). KLIPS 2차 웨이브의 가중치 부여방법에 관한 연구. 제2회 한국노동패널 학술대회 논문집.

강석훈. (2003). KLIPS의 가중치 부여방안 연구. 한국노동패널연구 2003-4, 한국노동연구원.

김규성, 박인호. (2010). 패널조사 웨이브 무응답의 대체방법 비교.

김규성, 황영은, 박진우. (2005). 패널조사에서 가중치 부여 방법 및 효과에 관한 연구. 제6회 한국노동패널학술대회 논문집.

김재광. (2017). 표본조사론. 자유아카데미.

김재광, 한근식, 윤연옥. (2004). 가계조사 무응답 처리기법 연구. 통계연구, 9(1), 79-102.

김태완, 이병재, 이주미, 최준영, 김기원, 박수지, 송치호, 주찬희. (2018). 2018년 한국복지패널 심층분석 – 해외 주요 패널과의 비교 연구. 한국보건사회연구원.

김현지, 김태용. (2018). 장애인패널조사 표본설계. 한국장애인개발원.

김현지, 김태용. (2019). 장애인삶 패널조사 가중치 산정 방안 연구. 한국장애인개발원.

김현지, 김태용, 이혜수. (2020). 장애인삶 패널조사 종단가중치 작성 방안 연구. 한국장애인개발원.

남기성, 천영민. (2012). An Analysis of Response Pattern and Panel Attrition in KLIPS. 응용통계연구, 25(6), 933-945.

박민규, 김사라. (2015). 한국노동패널조사 자료의 분석을 위한 패널 가중치 산출 및 사용방안 사례 연구. 조사연구, 16(1), 49-71.

박민규, 김영원, 변종석 (2013). 한국노동패널 가중치 연구. 한국노동연구원.

박민규, 이경상, 박현수, 강현철 (2011). 한국청소년패널조사(KYPS) 가중치 부여 방법 연구 : 중학교 2학년 패널의 경우. 조사연구, 12(3), 173-186.

박인호, 김기민, 김정우. (2020). 이중틀 접근을 통한 사업체패널조사의 표본추가 및 가중치 산출. 조사연구, 21(3), 37-59.

방하남, 장지연, 황수경, 김기헌, 김지경, 박시내.(2002). 한국노통패널 기초분석보고서(Ⅲ). 한국노동연구원.

성신여대 산학협력단. (2008). 패널자료 생성 및 이용활성화 방안 연구 – 농가경제조사(2003~2007)를 중심으로. 통계청 연구용역.

손창균 외. (2007). 자체통계 품질진단 매뉴얼 개발. 통계청 연구용역
손창균. (2011). 패널자료의 종단적 결측패턴에 관한 실증분석 연구. 한국조사연구학회 발표논문집, 서울: 한국조사연구학회, 273-275.
심규호, 백지선. (2012). 가계금융조사를 이용한 무응답 조정방법 연구. 2012년 하반기 연구보고서 제III권.
양수정. (2014). 장애인고용패널조사 유저가이드. 한국장애인고용개발원.
오미애, 이혜정, 이계오, 손창균, 박승환, 신재동. (2019). 2019년 한국복지패널 심층분석. 한국보건사회연구원.
윤윤규, 홍민기, 이사옳, 배기준, 최효미, 김기민, 장숙랑, 조성일, 천희란. (2012). 패널자료 품질개선 연구(2). 한국노동연구원.
이경상, 김기헌, 조주연, 김현희, 박영실. (2006). 한국 청소년 패널조사(KYPS) Ⅴ－조사개요 보고서-. 한국청소년정책연구원.
이경상, 김기헌, 임희진, 오소연, 소용준, 김유미. (2007). 한국 청소년 패널조사(KYPS) Ⅴ－조사개요 보고서-. 한국청소년정책연구원.
이경상, 안선영. (2009). 한국청소년패널조사VII : 1-6차년도 조사개요보고서. 한국청소년정책연구원.
이경상, 임희진, 안선영. (2008). 한국 청소년 패널조사 연구(KYPS). 한국청소년정책연구원.
이경희, 민인식. (2015). 패널조사 응답 지속성에 관한 연구-한국노동패널조사를 중심으로-. 조사연구, 17(3), 1-24.
이경희, 민인식. (2015). 패널조사 응답 지속성에 관한 연구: 한국노동패널조사를 중심으로. 한국노동연구원 KLI 패널 워킹페이퍼. 제2호.
이계오, 임찬수. (2012). 노동패널의 가중치 산출방안 : 원패널과 추가패널. 한국노동패널학술대회 논문집.
이규민, 강상진, 김현정. (2017). 한국교육종단연구의 표본설계, 자료의 질관리 및 분석방안 연구. 한국교육개발원.
이상호. (2005). 한국노동패널의 표본이탈 분석 -가구소득을 중심으로-. 노동리뷰. 11월호, 66-80.
이상호, 이혜정. 한국노동패널조사(KLIPS)의 패널이탈 분석-비단조 이탈을 중심으로.
이상호, 최효미. (2011). 다문화종단연구 패널연구방법론 Ⅲ: 패널조사의 데이터에러와 자료관리방안. 세미나자료집,
이상협, 박찬용, 정성석, 최혜미. (2011). 한국노동패널 탈락 분석. 한국데이터정보

과학회지, 22(1), 1-8.
이정화, 신슬비. (2017). 제4차 산재보험패널조사 기초분석 보고서. 근로복지공단 근로복지연구원.
이택면, 김승연, 우원규, 이주영, 강석훈. (2009). 2009년 여성가족패널조사 사업보고서 및 제2차 기초분석 보고서. 한국여성정책연구원.
이현우. (2009). 패널조사의 유용성과 문제점: 17대 대선자료를 대상으로. 조사연구, 10(2), 23-43.
이화정, 강석복. (2014). 한국아동·청소년패널조사자료에서 단위무응답의 실태 및 가중치 적용. 한국데이터정보과학회지, 25(6), 1397-1405.
전종우. (2008). 패널조사의 중요성과 발전방향. 보건복지포럼.
정미옥, 백지선. (2011). 가계금융조사를 위한 가구패널 설계 연구. 통계청 통계개발원.
정용찬, 유선실, 박민규. (2018). ICT 통계조사 품질진단. 정보통신정책연구원.
주재선, 김영택, 송치선, 손창균, 임찬수. (2012). 2012년 여성가족패널조사. 한국여성정책연구원.
최형아, 김영원. (2019). 패널조사에서 비연속 응답 그룹 편향 보정을 위한 복합가중값. 응용통계연구, 32(3), 375-389.
최효미, 성재민, 이상호, 배기준. (2012). 한국노동패널조사(KLIPS)의 CAPI 이행효과 : 실험설계 및 차수간 정합성을 중심으로. 조사연구, 13(3), 105-135.
최효미, 이상호, 성재민, 김기민, 배기준. (2012). 패널자료의 품질 개선을 위한 조사기법 연구: 컴퓨터를 이용한 조사(CAPI) 방법. 한국노동연구원.
하형석 외. (2018). 한국아동·청소년패널조사IX. 한국청소년정책연구원.
한국노동연구원. (2020). 한국노동패널 1-22차년도 조사자료 User's Guide.
홍민기, 김재광, 한치록, 김기민. (2014). 패널자료 품질개선 연구(III). 한국노동연구원.
홍민기, 이상호, 배기준. (2014). 조사방법과 히핑(heaping)과의 관계. 조사연구, 15(2), 123-147.
홍민기, 최효미. (2014). 노동패널 표본이탈과 불평등 지표. 조사연구, 15(4), 95-122.
Chang, W., Nishimura, R., Heeringa, S., Johnson, D., and Sastry N. (2021). panel study of income dynamics construction and evaluation of the 2019 longitudinal individual and family weights. survey research center, institute for social research university of michigan, Ann Arbor.

Elena Gouskova, Steven Heeringa, Katherine McGonagle, Robert Schoeni and Frank Stafford. (2008). Panel study of income Dynamics Revised Longitudinal Weights 1993-1995. Panel Study of Income Dynamics Technical Paper Series, University of Michigan.

Gouskova E., S. Heeringa, K. McGonagle, R. Schoeni, and F. Stafford. (2008). Panel Study of Income Dynamics Revised Longitudinal Weights 1993-2005, PSID Technical Series Report #08-05, ISR, University of Michigan.

Greg J. Duncan and Graham Kalton. (1987). Issues of Design and Analysis of Surveys Across Time. International Statistical Review, 55(1), 97-117.

John Fitzgerald, Peter Gottschalk, Robert Moffitt. (1996). AN ANALYSIS OF SAMPLE ATTRITION IN PANEL DATA: THE MICHIGAN PANEL STUDY OF INCOME DYNAMICS

Rainer Siegers, Veronika Belcheva, and Tobias Silbermann. 2020. SOEP-Core v35 – Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (SOEP) (1984 until 2018). SOEP Survey Papers 826: Series C. Berlin: DIW/SOEP

Robert Moffit, John Fitzgerald, and Peter Gottschalk. (1999). Sample Attrition in Panel Data: The Role of Selection on Observables. Annales d'Économie et de Statistique, Sep. - Dec., 1999, No. 55/56, Économétrie des Données de Panel / Panel Data Econometrics (Sep. - Dec., 1999), 129-152.

Lee, Una. (2003). Panel attrition in survey data: a literature review. CSSR Working Paper No. 41, 2-42.

Li Wang, Suojin Wang, Guannan Wang. (2014). Variable selection and estimation for longitudinal survey data. Journal of Multivariate Analysis, 130, 409-424.

Lynn, P. (2009). Methodology of Longitudinal Surveys, WILEY.

Mark Trappmann. (2015). The effect of events between waves on panel attrition. Survey Research Methods, 9(1), 31-43.

Nicole Watson. (2012). Longitudinal and Cross-sectional Weighting Methodology for the HILDA Survey, HILDA PROJECT TECHNICAL PAPER SERIES No. 2/12.

Peter Lynn. (2005). longitudinal surveys methodology. Willy.

Pudney, Stephen(2008). "Heaping and Leaping: Survey Response Behavior and the Dynamics of Self-reported

Uhrig, N. (2008). The Nature and Causes of Attrition in the British Household Panel Survey, ISER Working Paper Series, No. 2008-05.

Yiting Deng, D. Sunshine Hillygus, Jerome P. Reiter, Yajuan Si, and Siyu Zheng. (2012). Handling Attrition in Longitudinal Studies: The Case for Refreshment Samples.

한국장애인개발원
www.koddi.or.kr

장애인삶 패널조사 패널이탈 특성 및 종단가중치(3차년도) 연구

발 행 일 : 2021년 10월
발 행 인 : 최경숙
발 행 처 : 한국장애인개발원
서울시 영등포구 의사당대로22 이룸센터 5층
Tel. 02-3433-0600
Fax. 02-3433-9567
http://www.koddi.or.kr
편집·인쇄 : 누보

ISBN 978-89-6921-410-2(93330)

※ 본 연구보고서의 내용을 무단 복사하는 것을 금합니다.