

연구보고서 2018-39

2018년 한국복지패널 심층분석

- 해외 주요 패널과의 비교 연구



김태완 · 이병재 · 이주미 · 최준영 · 김기원 · 박수지 · 송치호 · 주찬희

【책임연구자】

김태완 한국보건사회연구원 연구위원

【주요 저서】

2017년 한국복지패널 기초분석보고서

한국보건사회연구원, 2017(공저)

2017년 기초생활보장 실태조사 및 평가연구

보건복지부·한국보건사회연구원, 2017(공저)

【공동연구진】

이병재 한국보건사회연구원 연구원

이주미 한국보건사회연구원 전문연구원

최준영 한국보건사회연구원 연구원

김기원 신성대학교 중어중문학과 교수

박수지 원주대학교 사회복지학과 교수

송치호 University of Washington 사회복지정책 박사

주찬희 서울시립대학교 사회복지학 박사

연구보고서 2018-39

**2018년 한국복지패널 심층분석
- 해외 주요 패널과의 비교 연구**

발 행 일 2018년 12월

저 자 김 태 완

발 행 인 조 흥 식

발 행 처 한국보건사회연구원

주 소 [30147]세종특별자치시 시청대로 370
세종국책연구단지 사회정책동(1층~5층)

전 화 대표전화: 044)287-8000

홈페이지 <http://www.kihasa.re.kr>

등 록 1994년 7월 1일 (제8-142호)

인 쇄 처 (사)아름다운사람들복지회

© 한국보건사회연구원 2018
ISBN 978-89-6827-563-0 93330

발간사 <<

한국복지패널은 국가승인 통계(제33109호)로, 2006년 조사가 진행되 이후 주요한 국가통계로서의 역할을 수행해 오고 있다. 한국복지패널의 연구진은 조사 결과의 품질 제고를 위해 각별한 노력을 기울였으며, 시의 성 높은 패널 데이터를 제공하고 학술적, 정책적으로 활용할 수 있도록 항상 노력하고 있다.

한국복지패널은 매년 기초분석보고서, 학술대회 보고서와 더불어 한국 복지패널의 방향성과 활용도 제고를 위해 별도의 심층분석 보고서를 발 간하고 있다. 2018년 심층분석 보고서에서는 한국복지패널을 해외 기관 및 연구자 등에게 알리기 위한 일환으로 첫째, 2012년에 진행된 패널 영 문화 사업을 계속해서 추진하였다. 둘째, 해외 패널에 대한 사례 연구를 통해 한국복지패널에 주는 시사점을 도출하였다. 셋째, 한국복지패널의 지속가능성 확보를 위한 방안으로 분가가구 활용 방안에 대한 기초연구 를 수행하였다. 올해 한국복지패널 심층분석 보고서에서 수행한 주요 연 구결과가 한국복지패널의 지속가능성과 외부 활용도를 높이는 데 기여하 기를 바란다.

본 연구는 우리 연구원의 김태완 연구위원의 책임하에 이병재 연구원, 이주미 전문연구원, 최준영 연구원이 참여하였다. 외부에서는 김기원 신 성대학교 교수, 박수지 원주대학교 교수, 송치호 University of Washington 사회복지정책 박사(수료), 주찬희 서울시립대학교 사회복 지학 박사 등이 함께 참여하였다.

본 연구가 한국복지패널의 개선과 활용도 제고 및 지속 발전을 위해 수

행된 보고서라는 점에서 사회복지, 빈곤, 공공부조 연구 등에 관심이 있는 교수 등 전문가 및 관련 분야 학생들에게 조금이나마 도움이 되기를 연구진은 바라고 있다.

2018년 12월

한국보건사회연구원 원장

조 흥 식

목 차

Abstract	1
요 약	3
제1장 서론	23
제1절 연구의 필요성 및 목적	25
제2절 연구 내용 및 방법	28
제2장 영국 British Household Panel Survey(BHPS)	31
제1절 조사 배경	33
제2절 표본 및 가중치	35
제3절 조사 방식 및 관리	40
제4절 조사표의 주요 내용	46
제5절 패널 활용	50
제6절 복지패널에 주는 시사점	53
제3장 미국 Panel Study of Income Dynamics(PSID)	55
제1절 조사 배경	57
제2절 표본 및 가중치	59
제3절 조사 방식 및 관리	69
제4절 조사표의 주요 내용	75
제5절 패널 활용	82
제6절 복지패널에 주는 시사점	84

제4장 독일 Socio Economic Panel(SOEP) 87

제1절 조사 배경	89
제2절 표본 및 가중치	92
제3절 조사 방식 및 관리	103
제4절 조사표의 주요 내용	114
제5절 패널 활용	122
제6절 복지패널에 주는 시사점	124

제5장 일본 Japan/Keio Household Panel Survey(JHPS/KHPS) .. 129

제1절 조사 배경	131
제2절 표본 및 가중치	134
제3절 조사 방식 및 관리	140
제4절 조사표의 주요 내용	147
제5절 패널 활용	154
제6절 복지패널에 주는 시사점	159

제6장 대만 Panel Study of Family Dynamics(PSFD) 161

제1절 조사 배경	163
제2절 표본 및 가중치	171
제3절 조사 방식 및 관리	177
제4절 조사표의 주요 내용	180
제5절 패널 활용	187
제6절 복지패널에 주는 시사점	191

제7장 결론	193
제1절 결론	195
제2절 한국복지패널에 주는 시사점	199
참고문헌	203
부록: 복지패널의 분가관리 방안을 위한 추적규칙과 가중값 조정에 대한 연구	209
제1절 패널 표본의 추적규칙	209
제2절 패널의 가중값 부여 방법	212
제3절 결론	226

표 목차

〈표 2-1〉 원표본 유지율	44
〈표 2-2〉 BHPS 가구용 조사표의 주요 내용	47
〈표 2-3〉 BHPS 개인용 조사표의 주요 내용	47
〈표 2-4〉 BHPS 부가조사 조사표의 주요 내용	49
〈표 4-1〉 독일 SOEP의 표본 특성 개관	96
〈표 4-2〉 표본 집단별 가구, 가구 구성원 조사 대상의 수	97
〈표 4-3〉 1984~2016년 표본 집단 규모(명)	98
〈표 4-4〉 인터뷰 조사원 투입 규모(2016년)	106
〈표 4-5〉 2016년 SOEP 조사에서 A~H 표본 집단의 조사 과정 결과	107
〈표 4-6〉 조사 집단별 인터뷰 방식의 비중(2016년)	110
〈표 4-7〉 PAPI와 CAPI에 평균적으로 소요되는 조사시간	112
〈표 4-8〉 가구 설문지 주요 문항	114
〈표 4-9〉 가구 구성원 개인 설문지 주요 문항	114
〈표 4-10〉 비독일인 응답자 주요 질문 문항	115
〈표 5-1〉 초기조사의 유의회수율(참고치)	142
〈표 5-2〉 KHPS 회답지속률	142
〈표 5-3〉 JHPS 회답지속률	143
〈표 5-4〉 한국과 일본의 패널조사 내용	149
〈표 5-5〉 International Seminar on KOWEPS & JHPS/KHPS	157
〈표 5-6〉 게이오대학 패널데이터 연구센터의 세미나·워크숍 일정(2018년도)	158
〈표 6-1〉 차수별 표본 추출 방법	172
〈표 6-2〉 표본 유실: 1999 표본(1953~1964)	173
〈표 6-3〉 표본 유실: 2000 표본(1935~1954)	173
〈표 6-4〉 표본 유실: 2003 표본(1964~1976)	174
〈표 6-5〉 표본 유실: 2009 표본(1977~1984)	174
〈표 6-6〉 표본 현황: 자녀표본	175

〈표 6-7〉 동남중국 조사-표본의 규모와 유실률 176

〈표 6-8〉 PSFD 연구 조직과 인력 179

〈표 6-9〉 PSFD 연구 자문위원 179

〈표 6-10〉 PSFD 공식회원 - 저널 & 콘퍼런스 논문 189

부표 목차

〈부표 1〉 BHPS에서 가중값의 종류 214

그림 목차

[그림 2-1] BHPS 데이터의 매칭(Matching BHPS data)	52
[그림 4-1] 2008~2016년 SOEP 표본 집단별 안정성	99
[그림 4-2] SOEP 패널의 데이터 수집 및 관리의 7단계	113
[그림 5-1] JHPS/KHPS의 4개 코호트	133
[그림 5-2] JHPS/KHPS의 구조	133
[그림 5-3] JHPS/KHPS의 표본 분포	136
[그림 5-4] KHPS2004 표본과 KHPS2007 추가 표본을 2007년 시점에서 통합	139
[그림 5-5] 게이오대학 패널데이터 연구센터 조직도	145
[그림 5-6] JHPS/KHPS 이용 실적	155
[그림 6-1] 조사 주기와 조사 대상	178

부도 목차

[부도 1] SLID의 패널 구성 형태	222
-----------------------------	-----

Abstract <<

The 2018 Korea Welfare Panel Study(KoWePS)

: An In-depth Report – International Comparison of Panel Studies in Selected Countries

Project Head: Kim, Taewan

In an attempt to broaden the use of The Korea Welfare Panel Study(KoWePS), the Korea Institute for Health and Social Affairs(KIHASA) has for some years published various reports and conference proceedings. This report, written in 2018, is an in-depth analysis of KoWePS. This report was begun with the intention of introducing the KoWePS and disseminating its outcomes to interested research institutions and researchers around the world.

This report explores the KoWePS from an international perspective, putting it in comparison with some of the major panel studies in other countries. We traced the processes through which these panels evolved over time and were made widely available and used in research and policy-making. This report focuses on some of the major long-standing panel surveys that have been conducted in the US, the UK, Europe and Asia. With the view to broadening the use of KoWePS, we had the Questionnaires and the User's guide translated into English.

Co-Researchers: Lee, Byeongjae · Lee, Jumi · Choi, junyoung · Kim, Giwon · Park, Susie · Song, Chiho · Joo, Chanhee

The panel surveys this study examined include the US' PSID, the UK's BHPS, Germany's SOEP, Japan's JHPS/KHPS and Taiwan's PSFD. In detail, we examined various aspects of the surveys, including their backgrounds, rationale, samples, sampling weights, questionnaires and data use, and drew policy implications for KoWePS. The authors hope that this report will be of use to those looking to improve KoWePS, which we hope will remain a rich source of authentic data for research and national policy-making.

1. 서론

- 2018년 복지패널은 13년 차 조사가 진행되었으며, 관련 자료의 축적과 홍보도 국내적으로 많이 이루어진 상태임.
 - 하지만 지금까지의 성과에도 불구하고 복지패널의 취약성으로 들 수 있는 부문이 해외 패널과의 교류라 볼 수 있음.
- 본 연구는 복지패널을 해외에 알리고 그 성과를 확산시킬 필요가 있다는 점에서 수행되었으며, 이와 같은 목적을 달성하기 위해서 두 가지 측면에서 연구과제를 진행함.
 - 첫째, 복지패널을 해외에 알리고 홍보하기 위해서는 먼저 해외 주요 패널의 역사적 발전 과정과 특성을 파악하고, 각국의 패널을 알리기 위한 노력이 어떻게 진행되고 있는지를 파악할 필요가 있음.
 - 둘째, 복지패널의 해외 연구자 활용도 제고를 위해 현재 국내 연구자에게 제공되고 있는 복지패널의 조사표, 유저가이드 등을 영문화하는 과정을 진행함.
- 본 연구에서는 미국과 영국, 유럽, 아시아에서 오랜 기간 조사가 진행되었고 활용도 또한 높은 주요 패널을 분석함.
 - 해외의 주요 패널로는 미국의 소득역동성패널연구(Panel Study

of Income Dynamics: PSID, 이하 PSID), 영국의 가구패널조사(British Household Panel Survey: BHPS, 이하 BHPS), 독일의 사회경제패널(Socio Economic Panel: SOEP, 이하 SOEP)를 선정하여 분석하였으며, 아시아 국가의 패널로는 일본의 가계패널조사(Japan/Keio Household Panel Survey: KHPS/JHPS, 이하 KHPS/JHPS), 대만의 가정 동태 데이터베이스(Panel Study of Family Dynamics: PSFD, 이하 PSFD)를 중심으로 연구를 수행함.

□ 본 보고서의 부록에는 분가가구 조사에 대한 검토 결과를 함께 수록하였음.

2. 주요 연구 결과

가. 영국 British Household Panel Survey(BHPS)

□ 조사 배경

○ BHPS는 에식스(Essex)대학 사회경제연구소(Institute for Social and Economic Research: ISER)와 경제사회조사협회(The Economic and Social Research Council: ESRC)의 영국 종단연구센터(UK Longitudinal Studies Centre: ULSC)가 수행함.

- BHPS는 1991년부터 시작되었으며, 영국 가정(개인)의 사회경제적 변화에 대한 이해, 변화의 원인과 결과 분석, 예측 등을 주요 목적으로 함.

- ULSC에서는 패널조사를 하고 연구에 활용할 수 있도록 제공하며 ISER은 BHPS 및 국외 패널을 활용하여 사회적 변화를 모니터링하고 측정하는 패널 데이터 기반의 연구를 수행함.

□ 표본 및 가중치

- 1차 웨이브(1991년)를 시작으로 대표성 확보를 위한 추가 표본에 대한 논의 및 수행이 지속되었으며 응답단위(개인, 가구)와 분석 목적(횡단면, 종단면 분석)에 따라 가중치를 제공하고 있음.
- 1991년 BHPS가 시작된 후 저소득층 표본(7차 웨이브), 스코틀랜드와 웨일스에 대한 표본(9차 웨이브)을 추가하였으며, 11차 조사부터는 북아일랜드 가구패널조사(NIHPS) 표본이 추가되면서 영국 전역의 표본을 확보하게 됨.
- 조사 대상은 초기 샘플 가구에 소속된 구성원(원표본멤버: OSM)과 추가적으로 원표본멤버(OSM)에게서 태어난 자녀, 원표본멤버(OSM) 중 한 명 이상이 포함된 분가가구의 가구원, 한 명 이상의 새로운 사람이 원표본멤버(OSM)로 편입된 가구원 모두 추가 조사 대상이 됨.
- 가중치의 경우 차수별, 응답자 개인과 가구별로 별도의 가중치를 제공하고 있으며, 분석 목적에 따라 차수별 횡단면 가중치 외에 종단면 가중치를 제공하고 있음.

□ 조사 방식 및 관리

- BHPS는 9차 웨이브부터 컴퓨터 지원 면대면 조사(Computer Assisted Personal Interview: CAPI, 이하 CAPI)를 도입하였으나 초기에는 서면 설문지 조사(Paper And Pencil Interview:

PAPI, 이하 PAPI)와 중복으로 조사를 하여 CAPI를 점진적으로 도입하였음.

- 또한 가구용과 개인용 설문에 대해서도 PAPI를 중복 실시함으로써 설문의 정확성을 높일 수 있는 이중의 안전장치를 마련하였음.
- 13차 웨이브부터는 대리용 설문에도 CAPI를 적용하기 시작하여 CAPI로의 이행을 종결하였음.
- 9차 웨이브의 현지조사 과정은 CAPI 도입 외에는 달라지지 않았는데, 이는 조사원의 경험과 교육이 CAPI로의 성공적인 전환을 위해 가장 중요한 요소였기 때문임.
- BHPS는 대면접촉을 기본 원칙으로 하지만, 불가능할 경우 대리면접을 허용하고 있으며 3차 웨이브부터는 전화조사도 일부 도입되었음.
- 개인정보가 포함된 문항에 대해서는 자기기입식 설문으로 조사를 수행하고 있음.

□ 조사표의 주요 내용

○ BHPS 조사표는 조사에 필요한 기본 정보가 담긴 표지(cover sheet) 외에 가구용(household), 개인용(individual)으로 나누어져 있으며 자기기입 조사표, 아동청소년 대상 조사표 등으로 구성되어 있음.

- 가구용은 주로 주택 유형과 주거 상태 등을 조사하며 부가적으로 가구의 생활을 유지하는 데 소요되는 소비지출 정보로 구성되어 있어 주거 상황에 대해 가장 많이 알고 있는 가구원을 대

상으로 함(백화종-김안나, 2004, p.84).

- 개인용 설문지에는 만 16세 이상의 모든 가구원이 응답하며 이웃과 개인의 인구학적 정보, 건강과 돌봄, 고용 상태, 직업력, 가치와 의견, 가구의 재무상태를 조사함(백화종-김안나, 2004, p.85).
- 부가조사에서는 자기기입 조사와 아동 조사 등을 수행하고 있음.

□ 패널 활용

○ 데이터 공개

- 학술적인 목적을 갖고 있는 모든 데이터 사용자는 데이터 보관소(UK Data Archive) 및 홈페이지에서 데이터에 접근할 수 있음. 데이터는 다양한 통계 및 분석 패키지와 함께 사용하기에 적합한 형식으로 제공됨.
- 데이터를 제공받기 위해서는 별도의 신청서를 제출해야 하며 센터에서 제공하는 양식에 서명을 해야 함. 데이터는 SPSS, SAS, STATA 데이터 파일로 제공됨.

○ BHPS 데이터 활용

- BHPS는 복지패널과 마찬가지로 15세 이상을 대상으로 하는 가구조사이며, 패널조사이기 때문에 매년 각 개인의 데이터를 축적하는 과정에서 표본의 이탈과 신규 추가가 발생하고 있음. 이에 따라 깔끔한 계층 구조가 형성되지 않고 비교적 복잡한 구조를 가지게 됨.
- 복잡한 구조의 데이터를 사용하기 위해서는 사용자가 여러 유형의 분석 단위로 재구성해야 함.

- 복지패널에서 가구키(h_merkey) 변수와 가구원키(h_pid) 변수를 식별변수로 활용하여 데이터를 재구조화하는 것과 같이, BHPS 또한 'KEYVAR'를 활용하여 개인을 색별하며 각각의 데이터에서 필요한 변수들을 연결할 수 있음.

○ BHPS 사용자 커뮤니티

- 센터는 데이터 사용자의 문제 해결과 정보 교환, 데이터 사용자 커뮤니티 구성을 위해 사용자 그룹을 관리하고 있음.
- 데이터 사용자 그룹과의 의사소통은 센터의 자체적인 연락망 외에 전자 뉴스레터 발송, 정기 뉴스 및 메시지 발송 등으로 하고 있음.

□ 복지패널에 주는 시사점

- 표본 유지와 관리는 국내와 유사한 방식으로 유지되고 있으며, 조사 방식에서도 성인을 대상으로 하는 조사와 아동을 대상으로 하는 조사가 구분된다는 점에서 유사한 특성이 발견됨.
- 연구센터에서 데이터 사용자 커뮤니티를 구성하고 직접 관리하는 것은 복지패널과의 차이점으로 볼 수 있음. 커뮤니티에서 지속적으로 데이터에 대한 의견을 정기적으로 제시하고 조사 내용 개발에도 참여할 수 있다는 것은 참고할 만한 부분임.

나. 미국 Panel Study of Income Dynamics(PSID)

□ PSID의 도입 배경과 특징

- 미국 36대 대통령 린든 존슨이 추진한 '빈곤과의 전쟁(War on

Poverty)’ 프로그램 평가를 위해 도입함.

- 미국 통계청은 미국 전체를 대표하는 약 3만 가구에 대한 조사를 완료함.
- 이후 경제기획국(OEO)은 가구의 경제적 복지에 대한 역동성을 파악할 목적으로 미시간대학 조사연구센터(the Survey Research Center, SRC)에 추가 조사를 의뢰함.
- 미시간대학 조사연구센터(SRC)가 진행한 이 조사연구가 PSID의 효시임.

○ PSID는 한국복지패널과 동일하게 저소득층을 과대표집한 대표적인 패널임.

- 1968년 시작된 이후 약 50년에 걸친 자료 축적.
- 조사 참여 대상자는 2017년 기준 약 7만 5천 명 이상, 최대 6세대 가구까지 포함.

○ PSID 자료는 다양한 학문분과(빈곤연구 및 심리학, 의학, 보건학 등)에 걸쳐 널리 활용되고 있음.

□ PSID의 표본과 가중치

○ PSID는 소득과 빈곤의 역동성 파악이라는 핵심목표를 달성하기 위해 1968년에 원표본을 구성하였고, 이 원표본은 2개의 하위 표본으로 구성됨.

- 미국 경제기획국(OEO)의 경제기획조사(SEO)에서 추출한 저소득층 1,872가구, 미시간대학 조사연구센터(SRC)에서 추출한 비빈곤층 2,930가구.

○ PSID의 표본은 차수에 따라 증가함.

○ PSID의 가족가중치는 기본적으로 ‘역(逆)의 확률 가중치 (inverse probability weights)’.

- PSID는 초기부터 가구 단위(household -level)가 아닌 가족 기준(family-level)으로 조사가 진행되며 가중치는 크게 종단면 그리고 횡단면 가중치로 구분됨.
- 종단면 가중치는 개인과 가족에 대한 2단계 작업을 거쳐 완성.
- 횡단면 가중치는 종단면 가족가중치에 근거하여 2단계를 거쳐 부여.

□ 조사 방식과 관리

○ PSID의 조사 방식은 크게 대면조사, 전화조사, 컴퓨터를 이용한 면접조사 방법 등 세 가지가 존재함.

- 조사 초기부터 1972년까지는 95% 이상이 면대면 조사로 수행되었고 이후 전화조사, 컴퓨터를 이용한 조사로 변화됨.

○ 1968년부터 1997년까지는 매년 조사되었고, 1997년 이후 격년으로 조사되고 있음.

- 핵심표본의 전체적인 응답률은 1968년 76%에서 2009년 94.3%로 꾸준히 증가함.

○ 조사 참여 대상자의 조사에 대한 거부감 완화 및 우호관계 증진을 위해 다양한 방법을 활용하며, 표본 마모를 방지하기 위한 다양한 인센티브를 제공함.

○ 조사의 체계적인 수행, 관리, 운영을 위해 미시간대학 조사연구센터(SRC)에 조사연구운영팀(Survey Research Operations, SRO unit)을 두고 있으며 미시간대학 교수진이 총괄적으로 지휘함.

- 1983년부터는 감독위원회(Board of Overseers)에서 조언을 받아 옴.

□ 주요 조사 내용

- 주요 조사 내용은 가구 일반 사항, 소득과 지출, 재산과 부채, 복지와 노동, 보건과 기타 영역의 다섯 가지를 중심으로 구성됨.
 - 소득과 지출은 조사 전년도 기준 세전소득을 조사하며 가족 전체 구성원들의 다양한 소득원천들의 합으로 구성.
 - 재산(net worth)은 설문지의 주거 영역과 부(또는 재산) 영역의 정보를 통해 계산하고 단기간 부채(short-term debts)를 포함한 아홉 가지 유형의 자산과 부채를 조사.
 - 그 외에 복지와 노동, 보건과 기타 영역을 조사하며 부가조사도 수행함.

□ 패널 활용

- PSID의 활용 실적은 네 가지 지표로 살펴볼 수 있음.
 - 첫째, PSID 자료를 이용한 출판 실적
 - 둘째, PSID 자료를 활용하는 연구에 대한 미국 국립보건원(NIH)과 국립과학재단(NSF)의 연구비 지원금(grants) 규모
 - 셋째, 등록된 사용자 수, 웹사이트 활동(방문 횟수), 자료 다운로드 횟수
 - 넷째, 사용자 학술대회 개최

□ 복지패널에 주는 시사점

- PSID는 국가 간 균등자료(Cross-National Equivalent File: CNEF, 이하 CNEF) 프로젝트를 수행하고 있음.
 - 한국복지패널도 PSID와 마찬가지로 국가 간 비교연구를 위한 자료구축 네트워크에 적극 참여할 필요가 있음.
- 사회구조적 요인의 변화와 그에 따른 복지욕구의 변화에 대한 반응성(responsiveness) 제고 차원에서 ‘패널의 조사 내용 다각화 및 타 자료와의 연계성 강화’가 필요함.

다. 독일 Socio Economic Panel(SOEP)

- 독일은 유럽에서 최초로 1984년 가구패널조사 SOEP를 시작함.
 - SOEP는 독일 인구 및 소득 통계의 수립을 목적으로 함. 고용, 교육, 자산, 소득, 개인이 가지고 있는 정부 정책에 대한 가치관 등을 매해 조사하고 있음.
 - SOEP는 1990년 6월부터 옛 동독지역 가구를 표본에 추가하여 조사하고 있으며, SOEP를 주관하는 공식 기관은 독일 주요 경제연구기관인 DIW Berlin(Deutsch Institut für Wirtschaftsforschung Berlin)임.
 - 독일 정부의 재정지원으로 수행되는 SOEP 패널은 유럽 각국에서 실시하는 가구 패널조사의 모태가 되었다는 평가를 받고 있음.
- SOEP는 해외의 주요 패널조사 중 가장 다양하고 복잡한 형태의 표본을 구성하고 있음.

- 1차 연도 조사 당시인 1984년 표본 A(통일 전 서독지역 독일인 표본), 표본 B(이탈리아, 스페인, 유고슬라비아, 터키 등 외국 국적을 가진 가구주로 구성된 표본) 5,921가구로 시작해 1990년도 표본 C(외국인은 제외하고 동독에 거주하고 있는 대상자를 중심으로 구성된 독일인 표본)를 추가함(송헌재, 2010, p.283).
- 표본 D는 1994·1995년도에 시작된 이민자 표본으로 표집 시기에 따라 D1과 D2로 구성됨(송헌재, 2010, p.283).
- 표본 E는 그동안 이민, 사망 등으로 탈락한 패널 가구를 보완하기 위해 1998년 독일 거주민 전체를 모집단으로 하여 추가한 것이며, 표본 F는 독일 정책의 효과성을 평가하기 위해 무자녀 가구, 복지수급자 가구 등을 모집단으로 하여 추가한 것임. 표본 G는 고소득자 표본이 과소표집되었다는 판단하에 소득 상위 2.5% 이내의 가구를 추가한 것임(송헌재, 2010, p.284).

〈요약표 1〉 독일 SOEP의 표본 특성 개관

(단위: 년, 가구, 명)

표본	시작 연도	가구	개인	표본
A, B	1984	5,921	16,205	서독 거주자 (A: 독일인 가구, B: 외국인 가구)
C	1990	2,179	6,131	동독에 거주하는 독일인 가구
D1	1994	236	733	1984년 이후 서독으로 이주한 가구
D1/D2	1995	522	1,665	
E	1998	1,067	2,470	서독에 거주하는 독일인 가구
F	2000	6,052	14,525	독일에 거주하는 독일인 및 이주민 가구
G	2002	1,224	3,538	월 소득이 7,500DM(3,835EURO) 이상인 가구
H	2006	1,506	2,616	독일에 거주하는 독일인 가구
J	2011	2,616	5,161	
K	2012	1,526		
L1	2010	1,247	2,311	독일 가족정책 평가 조사 가구 (아동 연령별 코호트 조사)

주: L2/3 표본은 스크린 표본으로 조사 설계가 상이해 여기에서 다루지 않음(CATI-CAWI-CAPI 모델).
자료: 이상호(2006), p.68; Anne B, et al(2018), pp.2-4.

□ SOEP의 2차 웨이브 이후 가구가중치와 가구원가중치는 다음과 같음.

○ 가구가중치에는 초기(1차)에 계산된 가중치를 부여함. 신규 가구원은 초기 가구가중치를 조정하는 방법을 사용하며, 이후 무응답 조정과 벤치마킹 조정을 실시함.

○ 가구원가중치는 가장 마지막에 형성된 가구가중치를 가구원에게 일괄적으로 부여함. 역시 마찬가지로 무응답 조정과 벤치마킹 조정을 실시함.

□ SOEP는 횡단 가중치 조정 방법으로 모형 방법(가구가중치 조정 → 가구원가중치)을 적용하고 있음.

□ SOEP는 대면면접을 통한 자기기입 방식, 즉 CAPI와 PAPI를 혼합

하여 사용하고 있음.

- 1996년 이전까지 SOEP는 다양한 표본으로 구성되어 각각의 표본에 따라 상이한 설문지가 설계되어 있었으나, 1996년 이후 표본 집단에 따른 설문지 차이가 없어지고 데이터 수집 또한 통합적으로 이뤄지고 있음.
- 표본 집단에 따른 설문지 차이는 없지만, 가구 구성원에 대한 측정 오차를 줄이기 위해 가구 설문지와 가구 구성원 개인 설문지를 구분하고 있음.
- 모든 설문 정보를 가구원 본인에게 직접 얻고자 하며, 조사원은 1인이 1구역을 맡아 관리함(이상호, 2006). 또한 매해 조사 과정에 대한 정보 또한 누적하여 메타 데이터 형식으로 저장하고 있음.

□ SOEP 설문지는 가구용, 가구원용으로 구분되어 있음.

- 조사표에 포함된 주요 내용은 인구학적 정보 외에 노동시장, 소득, 조세와 사회보장, 건강, 주거, 교육 및 직업훈련 등임.
- 2016년 SOEP 가구용 설문지의 문항은 다음과 같음.
 - 주거지와 주거비에 관한 질문
 - 정부보조금, 임대비, 이자소득 등 소득에 대한 질문
 - 대출납입금 혹은 적금 등 자산 형성을 고려한 수입에 관한 질문
 - 돌봄을 필요로 하는 동거 가구원, 돌봄을 필요로 하는 정도, 주 돌봄자에 관한 질문
 - 유치원, 학교 등에 재학 중인 아동에 관한 질문
- 2016년 가구원 개인에 대한 질문지가 대폭 수정되었는데, 질문

의 영역, 즉 설문 주제는 그대로 유지하되 현재 변화된 독일의 상황을 반영하여 개별 질문을 수정하였음.

- 전반적인 생활에 관한 질문
- 지난해의 경제활동 상황에 대한 질문
- 현재 가구원 개인의 경제활동에 대한 질문
- 건강과 질병에 대한 질문
- 주관적인 인식에 대한 질문
- 가족의 상황과 혈통에 대한 질문

□ SOEP의 데이터 관리는 설문조사를 총괄하는 TNS 인프라테스트(Infratest)에서 시행하고 있으며, 조사 방식에 대한 정보, 가구 정보, 개인정보, 해마다 누적된 정보 등을 분류하여 보관함.

○ SOEP 자료는 SIR-Export 파일, ASCII 파일, SAS/SPSS 파일로서 시디롬 형태로도 배포되며 다양한 활용을 지원하고 있음.

○ 또한 유럽 가구패널 컨소시엄(CHER: Consortium of Household Panels for European Socio-Economic Research)에 자료를 제공하고 있으며, CNEF로 활용되고 있음.

□ 복지패널에 주는 시사점은 다음과 같음.

○ 우리나라 또한 최근 이주민(다문화가정, 탈북자 등) 증가, 농어촌 인구 감소 등의 사회적 변화를 겪고 있으며, 이로 인한 인구 구조의 변화를 보다 심층적으로 고려한 조사 표본 설계를 위해 다양한 계획 및 전략이 필요함.

○ 또한 독일의 통일 대비 표본 설계 방식 계획, 사회 변화를 반영하

기 위한 지속적인 설문 문항 개선, 조사 과정의 기록을 의무화하여 조사 과정 전반에 대한 관리 시스템 구축, 국제적 교류 활성화 등의 결과는 향후 복지패널 연구에서 참고해 볼 만한 영역임.

라. 일본 Japan/Keio Household Panel(JHPS/KHPS)

□ 조사 배경

- 사회과학의 연구·정책평가 분석, 신뢰할 만한 데이터에 의한 평가 시스템 구축을 목적으로 2004년 게이오가계패널조사(KHPS), 2009년 일본가계패널조사(JHPS)를 실시함.
 - KHPS와 JHPS는 각각 개별 조사로 실시·관리되어 왔지만, 2014년부터 두 개의 조사표를 통일화하여 데이터 관리의 일원화를 도모하기로 함.
 - 따라서 일본을 대표하는 가계패널조사라는 의미를 담아 ‘일본가계패널조사(JHPS/KHPS)’로 명칭을 통일하기로 함.

□ 표본 및 가중치

- KHPS는 2004년 1차 조사 시 20세에서 69세까지의 남녀 약 4,000명을 대상으로, JHPS는 2009년 1차 조사 시 20세 이상(연령 상한 설정 없음)의 남녀 약 4,000명을 대상으로 함.
 - JHPS는 조사 시작 이래 신규 코호트가 없지만, KHPS는 2007년 1,400명과 2012년 1,000명을 신규 코호트로 추가함.
 - 데이터 표본·바이어스에 대한 수정이나 표본 탈락에 따른 바이어스 수정에 대한 대책은 조사 시작 이래 강구되어 오지 않았으나, 최근 게이오대학 패널데이터 연구센터에서 바이어스

에 대한 수정 가중치, KHPS와 JHPS를 통합할 때 필요한 표본 바이어스를 수정한 가중치를 발표함(石井 加代子, 野崎 華世, 2014).

□ 조사 방식 및 관리

- 연 1회, 1월 31일 시점에서 자기기입식으로 조사를 실시하며, KHPS는 전 대상자에게 자기기입식 방법으로 조사를 실시하지만 JHPS는 대상자에 따라 자기기입식 방법, 자기기입식과 면접(희망자는 인터넷상에서의 회답도 가능) 방법을 병용하여 조사함.
 - 대상자 관리방법으로는 대상자에게 인센티브로 사례 3,000 엔을 지급하고 있으며, 지속적으로 대상자와 연락하고 7월에는 복중문안(暑中見舞)엽서, 12월에는 뉴스레터를 보내고 있음.

□ 조사표의 주요 내용

- KHPS는 주로 취업, 소비, 소득, 주택 등을, JHPS는 취업, 소득, 교육, 건강·의료 등의 내용을 조사함.
 - 2014년 이후 조사 항목을 통일하여 가족 구성, 개인 속성, 학력, 취업·취학 상태, 생활시간 배분, 부모와의 동거관계 등 포괄적인 주제로 조사함.

□ 패널 활용

- 패널 데이터 자료 이용 및 연구 결과, 연구 성과 등을 홈페이지에 올려 다운로드하여 이용할 수 있게 함.
 - 게이오대학 패널데이터 연구센터는 비영리·학술 목적의 이용에 한해 국내외 대학 또는 국공립·민간 연구기관에 소속되어

있는 연구자·대학원생·학부생(연구생·연구원 등 포함)에게 데이터를 제공한다.

- 국제연계의 일환으로 룩셈부르크 소득연구원(Luxembourg Income Study: LIS, 이하 LIS)의 데이터베이스와 연계하고 있으며, CNEF와 연계하여 국가 간 비교가 가능한 패널에 포함될 예정임.

□ 복지패널에 주는 시사점

- 경제협력개발기구(OECD)의 소득변동 국제비교연구 등에서도 일본을 대표하는 데이터로 활용되는 등 국제적 활용도를 제고한다는 점에서 한국복지패널 또한 국제적으로 활용되기 위해 외부패널과의 협력이 필요할 것으로 보임.
- 또한 일본은 해외에서의 패널 활용을 파악하기 위해 외부에 패널 제공 시, 활용 논문에 대한 보고를 하도록 하고 있음. 한국 복지패널도 외부에서 얼마나 어떻게 활용되고 있는지를 파악하고 수치를 공개하여 복지패널의 위상을 알림과 동시에 활용도를 높이기 위한 방안을 마련해야 할 것임.

마. 대만 Panel Study of Family Dynamics(PSFD)

□ 조사 배경

- PSFD는 중국 가정의 성인표본을 대상으로 하는 고정표본 추적조사(panel)로 청소년부터 노인까지를 대상 연령층에 포함. 이 외에 본 계획은 주요 표본을 토대로 조사 대상을 자녀 등의 친족까지 확대하고, 다시 추적 방문조사를 함.

□ 표본 및 가중치

○ PSFD는 현재까지 대만에서 진행된 패널조사 중 기간이 가장 길고 방문 차수가 가장 많은 조사로 가정 중심(가족 구성원 중 한 명이 대표로 자신, 배우자, 부모, 자녀, 형제자매 등 가족관계에 응답) 조사임.

- 대만지구조사의 표본은 주표본과 자녀표본으로 구분할 수 있으며, 여기에서 자녀표본은 주표본의 16~24세 자녀가 해당함.
- 동남중국조사의 조사 지역은 상하이, 절강성, 푸젠성으로 대륙 5차 인구조사자료를 기반으로 비농업인구 점유 비중으로 단계를 분할한 각 단계에서 PPS법으로 표본을 추출함.

□ 조사 방식 및 관리

○ 대만지구조사는 2종의 표본(주표본, 자녀표본)과 3종의 조사(제1차 방문조사, 추적 주요조사, 자녀조사)로 이루어지며, 조사원의 방문을 통한 인터뷰 방식으로 조사가 이루어짐.

- 현지조사는 매년 1월 말에서 2월 초에 시작하여 약 1개월간 진행함. 매년 방문조사를 마친 후 피조사자에게 전화를 걸어 실제로 방문이 이루어졌는지 확인하는 절차를 거침.
- 조사 완료율을 높이기 위해 조사원이 50~150 TWD 상당의 답례품을 증정하거나 추첨행사, 연하카드, 조사 결과에 따른 보고서 등의 우편 발송 등을 함.

□ 조사표의 주요 내용

- 교육 경험과 직업 경험, 결혼과 배우자에 대한 내용, 결혼 여부, 부모 및 배우자의 부모, 형제자매에 대한 정보, 거주 형태, 가족 유대관계 및 지출(집안일 소요시간, 정부 보조금 수취 현황, 각종 지출 현황 등), 자녀교육과 관련한 내용이 포함됨.
- 소득 및 지출, 노동 등에 대한 세부 항목을 조사하고 있으나 복지와 관련한 세부 사항은 조사하고 있지 않음.

□ 패널 활용

- 조사에 대한 기본적 소개, 조사 활용 실적, 조사구조 및 자료 소개를 다루는 별도의 홈페이지를 운영하고 있으며, 홈페이지 회원에 한하여 자료를 제공함.

□ 복지패널에 주는 시사점

- PSFD의 경우 ‘중국인 가정의 경제, 사회, 심리, 관습 등에 대한 전반적인 연구’에 목적을 두고 있는 반면, 한국복지패널은 ‘한국 저소득층의 규모와 생활실태 파악’에 목적을 두고 있음.
- 따라서 PSFD 조사의 세부 사항이 ‘일반 중국 가정의 이해’에 해당하는 ‘효도와 가정에 대한 관념’, ‘가족 간의 친밀도’, ‘자녀 교육 태도’ 등이라면, 한국복지패널조사의 세부 사항은 ‘소득, 복지, 보건’ 등으로 볼 수 있음.
- ‘한국복지패널’에 ‘한국 저소득층 가정의 이해’와 관련된 세부 내역을 조사에 포함한다면, 현재 이루어지고 있는 경제적 정책 형성과 제도적 개선 외에도 저소득층의 가족관계 개선을 위한 구체적 사회복지제도 마련에 기여할 수 있을 것으로 기대됨.

3. 결론

□ 해외 패널과의 비교연구가 시사하는 점

- 첫째, 외부기관 등과의 협업으로 조사를 위한 펀딩 확보.
 - 장기적으로 데이터 설명회, 학술대회 혹은 별도의 콘퍼런스 등에 관련 기업이나 재단 등이 관심을 가질 수 있는 주제를 발굴하고 연구 결과가 발표된다면 연구재원을 확보할 수 있는 방안의 하나로 검토될 수 있음.
- 둘째, 패널 이용자를 관리하고 이용자를 대상으로 다양한 콘텐츠를 제공하며 패널을 활용할 수 있도록 교육프로그램을 주기적으로 제공.
- 셋째, 해외 패널과의 지속적 교류임. 영국, 독일, 일본, 미국 등의 주요 패널은 CNEF 혹은 자체적인 프로그램을 통해 해외 패널과의 지속적 교류를 진행.

□ 복지패널 중장기 표본 확보 방안

- 국내 대부분의 패널이 1세대 분가만을 허용하고 있다는 점에서 중장기적으로 1세대 분가를 넘어서 2세대 및 3세대로 넘어가는 분가가구에 대한 추적조사가 필요함.

*주요 용어: 복지패널 심층분석, 영국 BHPS, 미국 PSID, 독일 SOEP, 일본 JHPS/KHPS, 대만 PSFD, 분가가구

제 1 장

서론

제1절 연구의 필요성 및 목적

제2절 연구 내용 및 방법

제1절 연구의 필요성 및 목적

1. 연구의 필요성

- 한국복지패널(이하 복지패널)은 2006년 조사가 시작된 이후 지속적으로 가구의 경제 상태(소득·지출, 자산 등)뿐만 아니라 주거, 의료, 고용, 아동, 노인, 장애인, 가족, 복지인식 등 방대한 영역에 걸친 조사가 이루어지고 있음.
- 또한 복지패널 원자료를 활용하여 다양한 연구 분야의 주요한 쟁점을 분석함으로써 복지패널의 학문적·정책적 활용 가능성을 높이는 데 목적이 있음.
- 2018년 복지패널은 13차 조사가 진행되었으며, 관련 자료의 축적과 홍보도 국내적으로 많이 이루어진 상태임.
- 하지만 지금까지의 성과에도 불구하고 복지패널의 취약성으로 들 수 있는 부분이 해외 패널 변화에 대한 파악과 교류라 할 수 있음.
- 2017년 일본 게이오 패널과의 공동학술교류를 통해 우리 패널을 해외에 일부 알리는 성과를 거두었지만 여전히 해외 패널과의 교류(학문적 및 학술적)를 통해 복지패널의 지난 성과를 알리는 것 또한 중요한 패널의 역할이라 볼 수 있음.
- 한국복지패널이 해외 패널과의 교류를 확대하고 이를 통해 패널

개선 방향을 도출하기 위해서는 해외 패널의 운영, 조사 방법, 활용 방안 등에 대한 사전 연구가 필요함.¹⁾

□ 이 외에 복지패널조사가 진행되면서 발견된 중요한 장점 중 하나가 분가가구를 통한 세대 간 연구가 가능하다는 점임.

○ 패널조사가 장기간 지속됨에 따라 원표본 가구에서 분리된 분가가구가 지속해서 증가하고 있는 상황임.

○ 분가가구의 형태는 결혼, 취업, 군 입대 등 다양한 사유로 발생하고 있으며, 분가가구의 특성 연구는 자산의 이동, 빈곤의 대물림, 세대 간 연구에 많은 영향을 줄 수 있다는 점에서 중요함.

- 이 외에도 분가가구의 증가는 줄어드는 원표본가구를 보완해 주고 패널의 지속가능성을 높일 수 있다는 점에서 분가가구가 복지패널에서 매우 중요한 위치를 점유함.

○ 하지만 현재 복지패널에서는 분가가구의 범위를 원표본가구에서의 1세대 분가만을 허용하고 있음.

- 예를 들어 원표본 가구에 포함된 자녀가 결혼해 분가할 경우 해당 가구는 조사 대상에 포함. 하지만 원표본에서 분가한 가구의 자녀 혹은 특정 가구원이 재분가를 할 경우 그 가구는 조사에서 제외. 즉 분가가구의 분가를 허용하지 않고 있음.

○ 따라서 패널연구의 주요한 속성 중 하나인 세대 간 연구가 1세대에서 마무리되는 한계를 지님.

1) 원내의 논의 속에서 국내 패널 연구과의 교류에 대한 제안도 있었지만, 심층 논의 결과 복지패널의 내실을 다지기 위한 방안으로 해외 패널에 대한 조사와 교류를 먼저 시작하기로 함에 따라 동 연구를 진행함.

- 장기적으로 패널의 지속가능성과 패널조사가 가진 장점을 극대화하기 위해서는 현재와 같이 1세대 분가만을 조사할 것인지, 아니면 2세대 혹은 3세대까지 조사를 진행할 것인지에 대한 사전 검토가 필요함.

2. 연구 목적

- 본 연구는 한국복지패널을 해외에 알리고 그 성과를 확산시킬 필요가 있다는 점에서 수행되었으며, 위와 같은 목적을 달성하기 위해 두 가지 측면에서 연구과제가 진행됨.
 - 연구의 필요성에서 지적인 분가가구에 대한 논의는 향후 많은 검토와 토의 과정이 요구된다는 점에서 본 연구에서는 관련 전문가의 연구를 통해 나온 결과를 담는 것으로 같음하고자 함.
- 첫째, 복지패널의 향후 개선 방안을 도출하고, 우리 패널을 해외에 알리고 홍보하기 위해서는 먼저 해외 주요 패널의 역사적 발전 과정과 특성을 파악하고, 패널 개선을 위한 노력이 어떻게 진행되고 있는지를 파악할 필요가 있음.
 - 본 연구에서는 해외 주요 패널을 선정하여 주요 패널의 발전 과정, 특성, 주요 조사 내용 및 패널 관리 기법 등을 사전적으로 파악함.
 - 이를 통해 복지패널의 개선 사항을 도출하고 향후 해외 패널과의 교류를 위한 방안이 무엇인지 살펴보고자 했음.
- 둘째, 복지패널의 해외 연구자 활용도 제고를 위해 현재 국내 연구자에게 제공되고 있는 복지패널의 조사표, 유저가이드 등을 영문화

하는 과정을 진행.

- 2012년 일차적으로 조사표와 유저가이드 등에 대한 영문화가 진행된 바 있지만, 이후 이 영문화 자료에 대한 보완과 수정이 없었다는 점에서 6년이 지난 2018년을 기점으로 재영문화를 진행함.

제2절 연구 내용 및 방법

1. 연구 내용

- 해외의 주요 패널에 대한 분석은 우선 주요 패널을 선정하는 것에서 시작됨.
 - 본 연구에서는 미국과 영국, 유럽, 아시아에서 오랜 기간 조사가 진행되었고 활용도 또한 높은 주요 패널을 분석함.
 - 해외의 주요 패널로는 미국의 소득역동성패널연구(Panel Study of Income Dynamics: PSID, 이하 PSID), 영국의 가구패널조사(British Household Panel Survey: BHPS, 이하 BHPS), 독일의 사회경제패널(Socio Economic Panel: SOEP, 이하 SOEP)을 선정하여 분석하였으며,
 - 아시아 국가의 패널로는 일본의 가계패널조사(Japan/Keio Household Panel Survey: KHPS/JHPS, 이하 KHPS/JHPS), 대만의 가정 동태 데이터베이스(Panel Study of Family Dynamics: PSFD, 이하 PSFD)를 중심으로 연구를 수행함.
- 각각의 해외 패널에서 살펴보려는 주요 연구 내용에 대해 아래와 같

이 동일한 형태로 연구를 진행함.

○ 조사 배경: 조사 배경 및 주요 연혁

○ 표본 및 가중치

- 조사 대상(가구 수, 조사 대상 개인 수 등), 조사 대상(표본) 변화 및 가중치 부여 방식(횡단 및 종단, 원표본 및 통합가중치 등)
- 표본추출 방식, 무응답 처리 등

○ 조사 방식 및 관리

- 조사 방식: 면대면, CAPI, 조사 성공률 등
- 조사 주기(및 변화)
- 조직 및 인력(인력은 조사인력과 연구인력)

○ 조사표의 주요 내용

- 일반 사항: 조사 항목[조사가구 정의(가구 혹은 가족)]
- 소득 및 지출: 주요 항목(소득 및 지출의 정의, 가처분소득 계산 방식 등)
- 복지 및 노동, 보건 및 기타: 주요 항목
- 복지패널 조사표와의 비교

○ 패널 활용

- 홈페이지 운영, 외부 자료 제공 방식, 외부 패널과의 협력, 세미나 혹은 워크숍 개최, 활용 실적(논문, 보고서, 학위 등) 등

○ 복지패널에 주는 시사점

- 복지패널에 주는 시사점의 하나로 해외 주요 패널을 공유하고 있는 국가 간 균등자료(Cross-National Equivalent File:

CNEF, 이하 CNEF)와의 연계를 위해 각각의 패널이 어떠한 노력을 하고 있는지를 가능한 한 담고자 하였음.

□ 본 보고서의 부록에는 분가가구 조사와 가중치 조정에 대한 연구 결과를 수록하였음.

2. 연구 방법

□ 본 보고서의 주요 연구 내용을 위한 연구 방법은 해외 패널 연구에서 각국의 주요 패널에서 발간하는 자료와 홈페이지 등에 공개된 관련 정보를 활용함.

○ 추가적으로 필요시 관련 국내외 주요 문헌연구와 해외 패널에 대한 질문 등을 통해 해외 주요 패널에 대한 세부적 정보를 파악하고 관련 연구를 수행.

□ 복지패널의 조사표와 유저가이드 등에 대한 영문화 과제는 2012년의 일차 영문화 자료를 기초로 하여 지난 6년간 복지패널에서 추가(복지제도 변화 등)되거나 삭제된 정보를 사전에 파악.

○ 파악된 정보를 중심으로 일차 수정을 거쳐 영문화 과정을 수행하였으며, 이와 관련한 결과는 2018년 하반기에 국내외 연구자 등에게 복지패널 영문 홈페이지에서 공개.

□ 분가가구의 위상 제고와 향후 활용 방안은 관련 통계 전문가를 활용하여 연구하였으며, 최종 결과는 본 보고서의 부록에 수록하였음.

제 2 장

영국 British Household Panel Survey(BHPS)

제1절 조사 배경

제2절 표본 및 가중치

제3절 조사 방식 및 관리

제4절 조사표의 주요 내용

제5절 패널 활용

제6절 복지패널에 주는 시사점

2

영국 British Household << Panel Survey(BHPS)

제1절 조사 배경2)

□ BHPS는 에식스(Essex) 대학 사회경제연구소(Institute for Social and Economic Research: ISER, 이하 ISER)와 경제사회조사협회(The Economic and Social Research Council: ESRC, 이하 ESRC)의 영국 종단연구센터(UK Longitudinal Studies Centre: ULSC, 이하 ULSC)에 의해 수행되고 있음.

○ BHPS는 1991년부터 시작되었으며, 영국 가정(개인)의 사회경제적 변화에 대한 이해, 변화의 원인과 결과 분석, 예측 등을 주요 목적으로 하고 있음. BHPS는 광범위한 사회과학 분야의 연구 자원으로 설계되었으며 여러 분야에서 학제 간 연구를 지원함.

□ 영국 종단연구센터(ULSC)는 ESRC의 보조금을 지원받아 1999년 설립됨.

○ ULSC에서는 패널조사를 하고 이를 연구에 활용할 수 있도록 제공하고 있으며, ISER은 조사된 BHPS 및 기타 국외 패널을 활용하여 사회적 변화를 모니터링하고 측정하는 패널 데이터 기반의 연구를 수행함.

○ ULSC의 승인하에 2005년부터 ESRC에서 임명하는 과학운영위

2) BHPS의 조사 배경과 표본 추출에 대한 내용은 Taylor, M. et al(2010). BHPS User Manual 중 Introduction to the British Household Panel Survey(pp.23-33)를 참고하였음.

원회(Scientific Steering Committee)를 운영 중임.

- 이 위원회는 전국종단전략위원회(National Longitudinal Strategy Committee, NLSC)와의 협조 속에 BHPS에 대한 감독 및 관리 업무를 수행함.

□ BHPS는 1991년에 모집된 전국적인 대표성을 가지는 약 5,500가구와 영국 250개 지역의 1만 300명 이상의 개인(16세 이상)을 대상으로 매년 조사를 하고 있음.

○ 동일한 조사 대상을 추적하되 원래의 가정에서 분리되면 새로운 세대의 모든 성인도 조사 대상에 포함하며, 15세 이하의 아동은 16세가 되면 조사 대상에 포함함.

- 11~15세 아동에 대해서는 부가조사를 하고 있으며, 1999년과 2001년에 BHPS에 하위 표본이 추가되었음.
- 1999년의 경우 스코틀랜드와 웨일스에서 1,500가구가 추가로 표본에 들어왔으며, 2001년에는 북아일랜드의 2,000가구가 추가됨으로써 영국 전역에서 조사가 수행됨.

○ 2018년 기준으로 조사된 1~18차 웨이브를 외부에 공개하고 있음.

- BHPS 데이터는 영국데이터 서비스에 저장되어 있으며, 이를 사용하려는 경우 데이터 서비스에 등록하고 관련 라이선스를 받아 사용 가능.
- 사용하려는 정보에 따라 데이터 접근 방법에 차이를 두고 있으며, 세부적인 연구를 하려는 경우에는 별도의 라이선스를 받아야 데이터에 접근이 가능.
- 데이터는 SPSS, STATA 및 ASCII 등의 형식으로 제공되고 있음.

- BHPS 연구는 영국의 학술 및 정책 연구 커뮤니티에서 광범위한 협의를 거쳐 수립되었음.
- BHPS 연구의 주요 조사 항목은 가구 구성, 노동시장 참여, 소득 및 재산, 주거, 건강 및 사회경제적 가치 등이며
- 이를 통해 영국 가구의 건강 변화와 실업 간의 관계, 사회경제적 가치의 변화와 영향, 생애주기에 따른 소득의 변화, 노동시장에서의 수익과 지출, 주거 이동의 원인 및 결과 등을 분석하고 있음.

제2절 표본 및 가중치

1. 표본

- BHPS의 초기 샘플은 영국 우편번호 주소 파일에서 가져온 8,167 가구로 구성되었으며, 이 주소로 등록된 모든 가구에 조사가 시도되었음.
- 초기 표본 가구는 2단계 층화추출법과 체계적 표본추출 방법을 사용하여 선정됨.
- 표본추출에 사용된 틀은 영국 북부의 칼레도니아 운하(북아일랜드 제외)의 소규모 사용자용 우편번호 주소 파일(PAF: Postcode Address File)이며, 이것은 국가적으로 이루어지는 대규모 조사에서 일반적으로 사용되는 표본추출 틀임.
 - 초기 표본 추출은 효율성과 비용을 고려하여 2단계 층화집락 계통추출법을 활용함.

- 후속 조사에서도 초기 샘플로 선정된 모든 가구를 대상으로 조사를 하였음. 따라서 초기 샘플로 선정되었으나 조사를 거부한 가구에 대한 조사를 시도함. 계속해서 이어진 후속 조사에서는 거주지(영국, 웨일스, 스코틀랜드 등)에 따라 조사 대상 선정에 변화를 주었음.
- 초기 샘플 가구에 소속된 구성원을 원표본멤버(OSM: Original Sample Members)라고 지칭하며, 다음의 경우 추가 표본이 발생할 수 있음.
 - 원표본멤버(OSM)에게서 태어난 자녀
 - 원표본멤버(OSM) 중 한 명 이상이 새로운 가구로 편입
 - 한 명 이상의 새로운 사람이 원표본멤버(OSM)로 편입
- 원표본멤버(OSM)에게서 태어난 자녀는 자동적으로 OSM으로 간주됨. 또한 1년 중 6개월 이상을 OSM과 함께 거주하면서 생활을 공유하는 구성원 또한 조사 대상이 됨.
 - 1년 중 6개월 이상이라는 기준에 따라 방학 기간에 부모의 가정에 있는 학생은 조사 대상에서 제외되며, 학기 중에 거주하고 있는 가구의 구성원을 조사함.
 - 즉 차수별 표본은 모든 OSM과 임시 표본 멤버(TSM: Temporary Sample Members)로 구성됨.
- 1991년 BHPS가 시작된 이후 표본 추가는 여러 차례 진행됨.
 - 7차 조사부터 BHPS는 영국 유럽연합 가구패널(United Kingdom

European Community Household Panel: ECHP)에 데이터를 제공하기 시작하면서 북아일랜드의 모든 응답 대상과 영국패널(Great Britain panel)의 저소득층 표본을 추가함.

- 이때 저소득층 표본은 ECHP의 저소득층을 기반으로 선정됨.
- 따라서 목적에 따라 추가한 표본으로 1994년 ECHP 표본의 원표본 구성원은 BHPS의 OSM으로 정의하였으며, ECHP 2차 이후 응답 대상과 BHPS의 7차, 8차에 포함된 조사 대상은 BHPS의 표준 규칙³⁾에 따라 TSMs 또는 PSMs으로 정의됨.
- 2001년 ECHP가 중지됨에 따라 ECHP 예산을 더 이상 이용할 수 없게 되면서 11차 이후에서는 조사되지 않음.

○ BHPS의 9차 조사 시 주요 쟁점은 BHPS에 스코틀랜드와 웨일스에 대한 표본을 추가(또는 확대)하는 것이었음.

- 표본 확대에 대한 목적은 1) 상대적으로 작은 규모인 스코틀랜드와 웨일스 표본(초기 BHPS 시 각국별로 400~500가구 정도) 수를 추가하여 독립적인 분석이 가능하도록 하는 것, 2) 공공정책 변화의 영향을 평가하기 위해 영국과의 비교가 가능한 분석이 이루어질 수 있도록 하기 위함이었음.

○ 11차 조사부터는 북아일랜드 가구패널조사(NIHPS: Northern Ireland Household Panel Survey) 표본이 추가됨.

- 북아일랜드의 약 2,000가구를 추가하면서 그레이트 브리튼만이 아닌 영국 전역(UK-wide)의 패널을 확보하게 됨. 즉 전국

3) 1998년 표준 BHPS 규칙을 제정하였으며, 조사 대상 가구원에 대한 정의를 확인할 수 있음. 예를 들어 이혼, 독립 등의 사유로 분가한 가구의 경우 분가가구의 새로운 가구원은 표본으로 추가되어 모두 조사하며, 이때 기본적으로 1년 중 6개월 이상 연속으로 거주하는 경우 인정됨.

대표성을 확보하게 됨.

- 또한 이는 최근 영국과 북아일랜드를 비교할 수 있는 종단데이터를 확보하는 것이 북아일랜드의 정책입안자와 학계의 주 관심사였기 때문임.
- NIHPS의 첫 번째 조사 대상은 모두 OSM으로 BHPS의 표준 규칙이 적용됨.

2. 가중치

□ 가중치는 차수별, 응답자 개인과 가구별로 별도의 가중치를 제공하고 있으며, 분석 수준에 따라 사용할 수 있는 가중치는 달라짐.

○ 분석 목적에 따라 차수별 횡단면 가중치 외에 종단면 가중치가 제공되며, 종단 분석을 할 경우 마지막 차수의 종단면 가중치를 활용해야 함.

- 가중치는 가구가중치 외에 응답한 가구원과 기록된 가구원에 대한 가중치가 있으며, 이들 가중치는 주로 횡단면 분석에 사용됨.
- 기록가중치와 응답가구원가중치의 차이는 대리응답자, 전화 응답자 또는 16세 미만의 자녀 등은 기록개인가중치만 부여받는 방식이라는 점임.
- 종단면 분석에는 응답가구원가중치(wLRWGHT)와 기록가구원가중치(wLEWGHT)가 활용됨.

□ 1차 웨이브의 가중치는 다음과 같은 단계를 거쳐 가중치 조정이 이루어짐.⁴⁾

○ 불균등 선택 확률을 조정하기 위한 가중치(설계 가중치)

- 표본추출 과정에서 불균등한 추출확률을 조정하기 위해 1단계 조사구(Primary Sampling Unit: PSU)의 추출확률을 구한 다음 두 번째 단계로 1단계 PSU 상태에서 어떤 한 지점이 선택될 확률을 구함.
- 이러한 1, 2단계가 주어진 다음 선택된 어떤 지점에서 가구가 추출될 확률의 곱으로 추출확률을 계산한 후 이것의 역수로 가중치를 부여하는 방식.

○ 가구 단위에서 무응답 처리를 위해 조정하는 가중치

- 응답가구와 무응답가구 모두에 대한 가용 정보를 기반으로 지역, PSU의 특성, 주거 유형에 따라 조정됨.

○ 가구 내 개인의 무응답 처리를 위해 조정하는 가중치

- 응답가구 안에서 응답 대상 가구원임에도 무응답인 경우를 보완하기 위해 응답가구 내의 무응답가구원에 대한 조정이 필요함.
- 지역, 주거 점유 형태(보유 여부), 부유한 정도, 가구 내 응답 대상 가구원의 수, 혼인 여부, 고용 상태, 연령, 성별 및 상호작용 등의 변수를 적용하여 최적의 로짓모형을 설정하여 무응답 가구원에 대한 가중치를 조정함.

○ 최종 표본 크기에 따른 최종 가중치 재조정

- 가구가중치(AHHWGHT)[및 기록개인가중치(AXEWGHT)]는 설계가중치와 가구 단위의 무응답 가중치를 곱한 후 절삭

4) BHPS의 가중치 조정은 Taylor, M. et al(2010). BHPS User Manual 중 Weighting, Imputation and Sampling Errors(pp.184-195)', '손창균(2012). 패널 가중치 조정 방안에 관한 고찰. Technical Report. 패널포럼.' 자료의 내용을 참고하여 작성함.

(truncation)하고 모집단의 가구 수(및 개인 수)와 일치시키기 위해 사후층화한 다음 비율조정을 하여 표본 가구 수(및 표본 개인기록)와 일치하도록 함.

- 응답개인가중치(AXRWGHT)는 설계가중치에 가구 단위의 무응답가중치와 개인 단위의 무응답가중치를 곱한 후 절삭하고 모집단의 개입 총합과 일치시키기 위해 사후층화한 다음 비율조정으로 표본수와도 일치하도록 함.
- 모든 가중치 조정은 이전 차수에서 설정된 가중치를 기반으로 하여 실시되며, 따라서 이전 차수 가중치에 따라 결정됨.

□ 첫 번째 가중치는 2단계 층화추출설계로 발생한 편의를 조정하기 위해 적용됨. 표본 선정 과정은 확률표본추출 방법에 따라 설계됨.

○ 그러나 스코틀랜드와 잉글랜드, 웨일스에 분포된 조사 대상의 비중이 다르기 때문에 가중치 조정 작업이 필요함.

제3절 조사 방식 및 관리⁵⁾

□ ULSC와 현장 조사를 수행하도록 위탁받은 ‘국가여론조사(NOP: National Opinion Polls)’는 데이터 수집의 모든 측면에서 긴밀히 협력하여 효과적인 데이터 품질을 보장하기 위해 구조화된 조사를 수행함.

5) BHPS의 조사 방식과 관리 방법은 Taylor, M. et al(2010). BHPS User Manual 중 Sampling and Survey Methods(pp.133-182)를 참고하였음.

- ‘영국에서의 생활(Living in Britain)’이라고도 불리는 BHPS는 조사가 매우 복잡하고, 패널 표본을 지속적으로 유지해야 하기 때문에 ULSC와 조사 기관(NOP)은 매우 긴밀하고 지속적인 협력 관계를 유지하고 있음.
- NOP는 현장 조사, 편집, 코딩 및 데이터 입력을 체계적으로 관리하고 동시에 조사 설계에 관해 지속적으로 의견을 제시하고 있으며, 조사 지침 작성과 조사원 교육에 대해 일차적인 책임을 맡고 있음.
- 센터는 조사 현장의 의견, 코딩 요구사항을 수렴하여 데이터의 품질 관리에 중요한 역할을 하고 있음.
- ISER은 조사 설계, 설문지 개발 등 기술적인 측면의 책임을 지고 있으며, 실제 조사는 NOP(National Opinion Poll, Part of Market Research Ltd.)에 위탁해 이루어지고 있음.
- 초기에 조사원 배정은 250개 지역에 243명의 조사원으로 한 지역당 1인으로 구성하였는데, 이는 BHPS의 매우 복잡하고 난도가 있는 조사의 특성상 평균 이상의 경험과 능력이 필요했기 때문임.
- 조사원과 지역관리자는 다른 지역에서 열리는 총 14회의 브리핑 세션 중 한 번을 참석함.
- NOP와 연구센터 직원이 공동으로 주최하는 브리핑에 이전에 BHPS 경험이 있는 모든 조사원은 하루 참석하나, 신규 조사원의 경우에는 이틀 참석함.
- BHPS의 ‘영국에서의 생활(Living in Britain)’ 조사업무에 대한

이해를 높이기 위해 시청각자료 교육이 하루의 브리핑에 포함되어 있음.

□ 조사 기관에서는 표본 가구에 조사 목적을 간략하게 설명하는 우편물을 발송하며, 조사원은 가구를 방문하기 일주일 전에 전화로 연락함.

○ 모든 표본 가구는 조사 이후에도 지속적으로 조사에 대한 안내를 받으며 답례품을 받음. 표본 가구에 보낸 문서의 사본은 ISER에서 얻을 수 있음.

○ 표본 가구에서 조사를 거부할 경우 센터에서는 가능한 경우 추가 연락을 하고 특별서한을 보냄. 이 과정에서 상당수의 거부 가구를 다시 참여시킬 수 있음.

□ 조사는 매년 9월부터 12월까지 4개월에 걸쳐 이루어짐.

○ 대부분의 조사가 12월에 완료되지만, 다음해 3월까지 이사 가구의 추적과 조사 가구의 이의 신청에 대한 답변 등을 처리하는 작업이 수행됨.

□ 모든 조사원은 조사 초기에 감독자가 동반하여 면담 사본을 작성함. 감독자는 매주 두 번 현장에 나가 조사원의 조사 방식을 모니터링하고 조사진행률 등을 평가함.

○ 이 정보는 조사센터 데이터베이스로 전송·입력되며 주간 요약 보고서를 작성하는 데 이용됨.

○ 센터에서는 표본 가구에서 조사가 이루어졌는지 전화로 확인하고 개별조사 또한 수행되었는지 확인하고, 완료되지 않은 내용이나 추가 조사 또한 전화로 보완함.

- 조사품질 관리를 위해 조사 시간, 내용, 개인정보 보호, 민감한 정보에 대한 조사 등도 상시적으로 관리하고 있음.
 - BHPS는 대면접촉을 기본 원칙으로 실시하지만, 불가능할 경우 대리면접을 허용하고 있으며 3차 웨이브부터는 전화조사도 일부 도입되었음.
 - 개인정보가 포함된 문항에 대해서는 자기기입식 설문으로 수행하고 있음.
- BHPS는 가구의 분가를 추적해서 이루어지는 패널조사인데, 개인이 원가구에서 분리되지 않는다면 다음 연차 조사에서 다시 면접함.
 - 가구에 새로 진입한 가구원이 성인이라면 조사 대상이 됨. 즉 조사 연도 조사가구에 포함된 모든 가구원을 대상으로 조사를 수행함.
 - 결국 표본은 출생이나 혼인, 동거 등으로 새로운 가구원이 늘어나기도 하고 사망, 조사 거부, 이사 등으로 감소하기도 함.

〈표 2-1〉 원표본 유지율

(단위: 가구, %)

구분	2차		9차		18차	
조사 완료	8,568	86.4	6,396	64.5	4,411	44.5
대리 응답	150	1.5	85	0.9	36	0.4
전화 응답	-	-	47	0.5	270	2.7
일부 가구원 응답 거부	112	1.1	135	1.6	155	1.6
일부 가구원 무응답(기타 이유)	12	0.1	19	0.2	27	0.3
사망	81	0.8	771	7.8	1570	15.8
거주지 이탈	64	0.6	269	2.7	464	4.7
부재(추적 불가)	267	2.7	286	2.9	223	2.2
가구 전체 응답 거부	620	6.2	206	2.1	143	1.4
가구 전체 무응답(기타 이유)	38	0.4	129	1.3	156	1.6
이전 차수부터 응답 거부	-	-	1,235	12.7	1,942	19.6
장기 부재	-	-	294	3.0	515	5.2

자료: Taylor, M. et al(2010). BHPS User Manual, pp.180-181.

□ 9차 웨이브에서 데이터 수집방식을 서면 설문지 조사(Paper And Pencil Interview: PAPI, 이하 PAPI)에서 컴퓨터 지원 면대면 조사(Computer Assisted Personal Interview: CAPI, 이하 CAPI)로 전환하였음.

○ CAPI로 전환한 주요 원인은 CAPI시스템이 데이터 질 향상에서 강점을 가지고 있으며, BHPS는 개인용 조사에 45분이 소요되고, 복잡한 로직을 가지고 있어 조사원으로부터 발생할 수 있는 오류들을 일관되고 정확하게 입력되도록 할 수 있음.

○ 응답의 결측치를 낮추기 위한 방안으로 응답 회피 항목이 존재하지만 반드시 응답해야만 다음 질문으로 넘어가도록 설계하였음.

- BHPS는 9차 웨이브부터 CAPI시스템을 도입하였으나 초기에는 PAPI와 중복으로 조사하여 CAPI의 점진적 도입을 실시하였음.
- 또한 가구용과 개인용 설문에 대해서도 PAPI를 중복으로 실시함으로써 설문의 정확성을 높일 수 있는 이중의 안전장치를 마련하였음.
- 13차 웨이브부터는 대리용 설문에도 CAPI를 적용하기 시작하여 CAPI로의 이행을 종결하였음.
- 9차 웨이브의 현지조사 과정은 CAPI 도입 외에는 달라지지 않았는데, 이는 조사원의 경험과 교육이 CAPI로의 성공적인 전환을 위해 가장 중요한 요소였기 때문임.
 - 응답자에 대해 동일한 조사원을 유지하는 것이 설문조사 기간 응답률에 도움이 된다는 선행연구(Laurie, H. et al, 1999, Taylor, M. et al, 2010 재인용)와 같이 응답자와의 지속성과 라포르 유지를 위해 이전 조사를 진행한 가구에 동일한 조사원을 배치하는 정책을 가지고 있음.
 - 정규 BHPS 조사원의 다수는 현지에서 NOP의 CAPI시스템을 사용한 경험이 있었고, CAPI 도입으로 탈락한 정규 조사원은 없었음. 주요 설문조사에 참여한 220명의 조사원 중 16명만이 9차 웨이브의 BHPS에서 새로 채용되었으며, 모두가 이전에 CAPI의 경험을 가졌음.

제4절 조사표의 주요 내용

□ 조사표 구성

○ BHPS 조사표는 조사에 필요한 기본 정보가 담긴 표지(cover sheet) 외에 가구용(household), 개인용(individual)으로 나누어져 있으며 자기기입 조사표, 아동청소년 대상 조사표 등으로 구성되어 있음.

- 가구용은 주로 주택 유형과 주거 상태 등을 조사하며 부가적으로 가구의 생활을 유지하는 데 소요되는 소비지출 정보로 구성되어 있어 주거 상황에 대해 가장 많이 알고 있는 가구원을 대상으로 함(백화종·김안나, 2004, p.84).
- 개인용 설문지에는 만 16세 이상의 모든 가구원이 응답하며 개인 및 이웃에 대한 인구학적 정보, 건강과 돌봄, 고용 상태와 직업력, 소비와 지출 등 매우 광범위한 범위로 구성되어 있음(백화종·김안나, 2004, p.85).
- 부가조사에서는 자기기입 조사와 아동 조사 등을 수행하고 있음.

□ 각 조사표에서 담고 있는 주요 내용은 다음과 같음.

○ 조사 기본 정보(cover sheet)

- 성별, 연령, 지역(잉글랜드, 웨일스, 스코틀랜드), 가구원의 분가(이혼, 진학, 구직 등) 또는 합가(결혼, 출산 등) 유형 등

○ 가구용(household)

- 가구용 조사표의 주요 내용은 <표 2-2>와 같음.

〈표 2-2〉 BHPS 가구용 조사표의 주요 내용

영역	내용
주거	- 주거 유형(단독주택, 연립주택, 아파트, 사업장 거주, 보호시설 등) - 자가/임대 여부, 방 개수, 주택 가격, 주택 구입 또는 임대 방법, - 주택담보대출 여부, 대출기간, 대출금액, 대출액을 제외한 추가 용자 여부와 용자 금액, 월 납입금액 등 - 상하수도 요금, 난방비, 수리 및 유지 보수 비용 등 - 주거 임대 시 받은 혜택(공제 등) - 주거비로 인한 어려움의 유무, 어려운 정도와 해결 방안 - 주택 내부 편의시설 유무 - 난방 장치 유무, 연료의 종류, 연료비 사용 정도 - 태양력, 풍력 등 친환경에너지 사용 여부, 시설 보유 현황 - 쓰레기 배출, 재활용쓰레기 분리수거 수준 - 주거 관련 각종 어려움: 공간 부족, 소음, 채광 등 환경문제, 난방 설비, 누수, 기타 부실공사, 교통, 지역 범죄 등 - 주거 중인 주택 외의 부동산 소유 여부와 가격
소비와 지출	- 가전제품 소유 현황과 구입 가격 - 인터넷 사용 여부 - 부동산 외 대출 여부, 경제적 부담 정도 - 난방비, 휴가비, 의복비, 식료품비, 문화·여가비 등의 지출 여부와 지출 가능한 수준 - 차량 소유 여부, 차량 구매 시 부채 여부

자료: BHPS Questionnaires (Wave 18), pp.24-45.

○ 개인용(individual)

- 개인용 조사표의 주요 내용은 〈표 2-3〉과 같음.

〈표 2-3〉 BHPS 개인용 조사표의 주요 내용

영역	내용
이웃과 개인의 인구학적 정보	- 출생 연도와 출생지, 주거 기간, 현재 시민권 상태, 소속된 국가(잉글랜드, 스코틀랜드, 웨일스), 국가별 언어 사용의 차이 정도, 국가별 이주 현황 - 이사 여부와 이사 이유 - 부모의 직업과 출생지, 부모의 경제활동 상태, 조부모의 출생지 - 본인 경제활동 상태 - 교육 참여 경험, 교육 수준, 현재 교육 상태, 참여한 교육 과정의 기간과 교육 내용, 교육 참여 시 연령 - 직장 내 교육제도 활용 여부 - 직업 능력 관련 교육 이수 여부, 학업, 직업 관련 기술 또는 자격 취득 여

영역	내용
	- 부, 취득한 기술 및 자격의 종류 - 현재 혼인 상태와 변화 여부(사실혼 포함) - 출산 여부, 태어난 아기의 기본 정보, 건강 상태 - 가정과 이웃, 지역사회에 대한 만족도 - 인터넷 사용 여부 - 비혼 상태에서의 파트너 유무, 향후 결혼(재혼) 가능성 또는 비혼 가능성, 결혼이 아닌 동거 또는 교제만을 선택한 이유, 장점과 단점, 만족도
건강과 돌봄	- 장애 여부 - 주관적·객관적 건강 상태 - 일상생활 유지 정도 - 일상생활 유지 정도: 걷기, 계단 오르기, 집안 관리, 위생 관리 등 - 병원 방문 빈도, 병원에서 보낸 시간 - 입원, 사고 경험 - 진료비 지불 방법: NHS, 개인 등 - 민간 의료보험 가입 여부, 보험료 - 의료복지서비스 이용 경험, 이용 빈도, 이용료 - 건강검진 경험, 검진 분야, 횟수 - 흡연 여부, 흡연량 - 돌봄이 필요한 가구원의 유무, 관계, 돌봄 제공 주체, 돌봄에 사용하는 시간
고용 상태	- 주된 일자리 유형, 유급노동 여부, 풀타임/파트타임 여부, 직급, 파트타임 일자리 종사 이유, 일자리 유지 기간 - 회사 유형, 직업군, 고용주 또는 고용인과의 관계, 종사자 수, 노동 시간 - 출퇴근 소요시간, 교통수단 - 직업 만족도 - 임금(세전/세후), 근로세액공제(Working Tax Credit) 등 각종 공제 혜택 여부 - 시간당 임금, 초과근무 빈도·시간·수당 - 성과급 유무, 급여 인상 수준, 노동조합 등의 임금협상 가능 여부, 조직 내 승진기회 제공 여부, 조직이 운영하는 연금제도 유무와 가입 여부, 유연근무제 등 시행 등 종사자 처우 수준 - 향후 이직/퇴직 계획, 직업훈련 경험 - 근무 중 자녀 돌봄 방법, 직장 내 보육시설 유무, 보육 비용 - 구직 센터 이용 경험
직업력	- 현재의 경제활동 상태(임금근로, 자영업, 은퇴, 교육훈련, 휴가, 학생 등) - 과거의 직업 이력과 당시 구체적인 업무 내용 - 당시 고용지위, 직급, 고용 기간, 회사 규모, 종사자 처우 등 - 이직·퇴직을 결심한 이유, 향후 구직 가능성
가치와 의견	- 국가와 정부에 대한 만족도 - 정치에 대한 관심도, 지지하는 정당의 유무와 지지하는 정도, 정치이념, 노동조합 또는 이익집단과의 소속 여부 - 종교 유무, 집회 참석 빈도 - 여가생활의 유형 - 친구, 친척, 이웃과의 관계, 왕래 빈도, 관계에 대한 만족도

영역	내용
	- 건강, 재산, 가족, 직업 등을 중요하게 여기는 정도 - 교육, 의료, 교통 등 지역사회 서비스와 인프라에 대한 만족도 - 국외 여행이나 방문 경험과 빈도 - 환경 보호에 대한 관심 수준
가구 재무상태	○ 소득 - 각종 연금급여(공공, 민간, 퇴직, 크레딧 등) - 장애, 상병, 보훈 등의 사유로 받은 급여 - 근로장려, 보육 및 교육, 주거, 구직 등의 사유로 받은 소득 보전 - 임대, 보험, 저축, 조합 활동 등으로 인한 각종 수입 ○ 지출 - 생활비(식료품, 의복, 통신 등)와 소요 시간 - 여가활동비와 여가활동시간 - 사적 이전(주거, 교육, 대출상환 등) - 양육비와 양육시간 - 차량 소유와 관리

자료: BHPS Questionnaires (Wave 18), pp.46-183.

○ 부가조사

- 부가조사는 자기기입 조사와 아동 조사 등이 있으며 주요 내용은 <표 2-4>와 같음.

<표 2-4> BHPS 부가조사 조사표의 주요 내용

영역	내용
자기기입 (Self-completion)	- 걱정과 불안, 우울에 대한 인지 - 사물과 현상에 대한 인지 - 긴장 상태와 어려움에 직면하고 대처하는 방식 - 사회 정의, 질서, 규율, 환경, 생명에 대한 가치
아동 (Child)	- 시간 사용 - 가족, 친구, 사회적 관계 진단 - 건강과 영양 - 정신건강 - 이념과 가치 - 향후 학업, 직업 선택 등에 관한 계획

자료: BHPS Questionnaires (Wave 18), pp.210-318.

제5절 패널 활용⁶⁾

□ 데이터 공개

- 학술적인 목적을 갖고 있는 모든 데이터 사용자는 데이터 보관소 (UK Data Archive)를 통해 데이터에 접근할 수 있음. 데이터는 다양한 통계 및 분석 패키지와 함께 사용하기에 적합한 형식으로 제공됨.
 - 데이터에는 BHPS 홈페이지(<https://www.iser.essex.ac.uk/bhps>)에서 자유롭게 접근할 수 있음.
- 상업적인 목적의 분석을 수행하는 연구자는 데이터를 이용할 때 연구센터에서 비용을 청구할 수 있음. 비용 청구에 대한 결정은 연구센터에서 하고 데이터 보관소를 통해 관리함.
- 데이터를 제공받기 위해서는 별도의 신청서를 제출해야 하며 센터에서 제공하는 양식에 서명해야 함.
- 연구센터는 조사 결과를 SPSS, SAS, STATA 데이터 파일로 구성하여 데이터 아카이브를 통해 제공함.

□ 데이터 활용

- ESRC 연구센터(ESRC Research Center)는 패널데이터의 사용자 커뮤니티에서 데이터 사용을 촉진하는 데 관심을 기울이고 있음.
- 센터는 사용자가 데이터를 사용하던 도중 발견하게 되는 오류나

6) BHPS 데이터의 구조와 활용에 대한 내용은 Taylor, M. et al(2010). BHPS User Manual 중 The BHPS Data(pp.36-43), Data Dissemination(pp.214-216)을 참고 하였음.

불일치 정보는 센터에 알려주도록 요청하고 있으며, 수정사항 또한 모든 BHPS 데이터 사용자가 즉시 알 수 있도록 통보하고 있음.

- 차수별 데이터는 다음해 데이터와 함께 다시 제공되기 때문에 사용자와 센터 간에 정보를 주고받는 것은 특히 중요한 작업임.

□ BHPS 데이터 구조

○ BHPS는 복지패널과 마찬가지로 15세 이상을 대상으로 하는 가구 조사이며, 패널조사이기 때문에 매년 각 개인의 데이터를 추적하는 과정에서 표본 이탈과 신규 추가가 발생하고 있음.

- 다양한 수준(가구 및 개인)에서 수집되는 이러한 데이터는 시간이 지남에 따라 깔끔한 계층 구조로 형성되지 않고 비교적 복잡한 구조를 가지게 됨.

○ 복잡한 구조의 데이터를 사용하기 위해서는 사용자가 여러 유형의 분석 단위로 재구성해야 함.

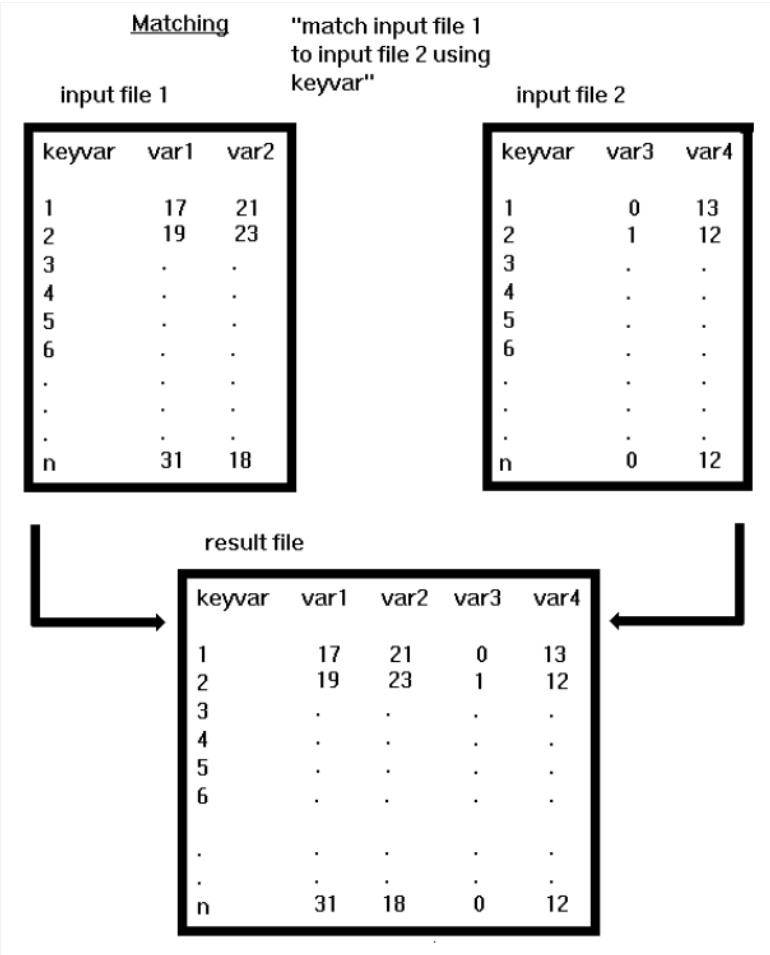
- 가구와 가구의 기본 정보, 직업, 수입과 지출, 사회적 관계 등이 각각 담겨 있는 여러 단위의 데이터를 연결하면 분석 범위가 넓어지고 다양한 분석을 가능하게 하므로 사용자가 원하는 방식으로 데이터를 구성해야 함(data merge).

○ 복지패널에서 가구키(h_merkey) 변수와 가구원키(h_pid) 변수를 식별변수로 활용하여 데이터를 재구조화하는 것과 같이, BHPS 또한 'KEYVAR'를 활용하여 개인을 식별하며 각각의 데이터에서 필요한 변수들을 연결할 수 있음.

- [그림 2-1]은 이러한 절차를 간단하게 보여 주고 있음. 'file 1'에서는 변수 'var1'과 'var2'가 필요하고 'file 2'에서는 변수

‘var3’과 ‘var4’가 필요한 경우, 두 개의 서로 다른 데이터가 포함된 정보를 ‘keyvar’를 기준으로 연결하여 각 변수들의 관계를 분석할 수 있음.

[그림 2-1] BHPS 데이터의 매칭(Matching BHPS data)



자료: Taylor, M. et al(2010). BHPS User Manual, p. 37.

□ BHPS 사용자 커뮤니티

- 센터는 데이터 사용자의 문제 해결과 정보 교환, 데이터 사용자 커뮤니티 구성을 위해 사용자 그룹을 관리하고 있음.
 - BHPS는 종단 패널 연구이므로 향후 데이터 수집을 계속하면서 사용자와 정기적인 접촉을 유지하는 것은 센터의 자산이 됨. 사용자는 향후 데이터 수집에 의견을 제시하고 조사 내용 개발에 참여할 기회를 얻을 수 있음.
- 데이터 사용자 그룹과의 의사소통은 센터의 자체적인 연락망 외에 전자 뉴스레터 발송, 정기 뉴스 및 메시지 발송 등으로 하고 있음.
 - 연구센터는 광범위한 전문 지식을 가진 사용자 그룹을 관리함으로써 데이터의 발전과 지식의 교류 등으로 전체 사회과학 커뮤니티를 풍성하게 만들 수 있을 것으로 기대함.

제6절 복지패널에 주는 시사점

- 표본 유지와 관리는 국내와 유사한 방식으로 유지되고 있으며, 조사 방식에서도 성인을 대상으로 하는 조사와 아동을 대상으로 하는 조사가 구분된다는 점에서 유사한 특성이 발견됨.
- 조사 내용 또한 복지패널과 유사하게 가구 구성, 노동시장 참여, 소득 및 재산, 주거, 건강 등의 영역을 조사하고 있음.
- 연구센터에서 데이터 사용자 커뮤니티를 구성하고 직접 관리하고 있는 것은 복지패널과의 차이점으로 볼 수 있음. 커뮤니티에서 지속

적으로 데이터에 대한 의견을 정기적으로 제시하고 조사 내용 개발에도 참여할 수 있다는 것은 참고할 만한 부분임.

제 3 장

미국 Panel Study of Income Dynamics(PSID)

제1절 조사 배경

제2절 표본 및 가중치

제3절 조사 방식 및 관리

제4절 조사표의 주요 내용

제5절 패널 활용

제6절 복지패널에 주는 시사점

3

미국 Panel Study of Income Dynamics(PSID) <<

제1절 조사 배경⁷⁾

□ PSID의 도입 배경

- 미국 36대 대통령 린든 존슨이 추진한 ‘빈곤과의 전쟁(War on Poverty)’ 프로그램 평가를 위해 도입.
 - 미국 경제기회국(the Office of Economic Opportunity, OEO)은 경제기회조사(the Survey of Economic Opportunity, SEO)를 1966년과 1967년에 미국 통계청(the US Census Bureau)에 의뢰.
 - 미국통계청은 미국 전체를 대표하는 약 2만 2,000가구와 비백인(non-white) 인구를 대표하는 약 1만 5,000가구를 경제기회조사(SEO)의 표본가구로 설정.
 - 경제기회조사(SEO) 결과, 약 3만 가구에 대해 조사 완료.
 - 경제기회국(OEO)은 가구의 경제적 복지에 대한 역동성을 파악할 목적으로 경제기회조사(SEO) 표본 중 저소득가구로만 구성된 하위 표본 조사를 지속하길 기대.
 - 이를 위해 경제기회국(OEO)은 미시간대학 조사연구센터(The Survey Research Center, SRC)에 해당 저소득 표본(소득 하위 2천여 가구) 조사를 의뢰.

7) 이하의 내용은 McGonagle, K. et al.(2012)를 참고하였다.

- 하지만 당시 미시간대학 연구책임자였던 제임스 모건(James Morgan)은 미국 전체 인구의 대표성 확보를 위해 경제기획조사(SEO)의 저소득가구표본에 미시간대학 조사연구센터(SRC)의 비빈곤(non-poor) 가구표본을 포함해야 함을 역설.
 - 경제기획국(OEO)은 모건의 주장을 결국 수용.
- 원표본가구에서 분리된 가구원(예를 들어 성년이 되어 부모에게서 분가한 자녀)을 추적하여 조사토록 한 우연한 결정이 존재.
- 이러한 결정 때문에 해당 표본은 시계열적으로 미국 전체 개인과 가구 대표성 유지가 가능.
- 결국 미시간대학 조사연구센터(SRC)가 진행한 이 조사연구가 PSID의 효시.

□ PSID의 특징

- PSID는 한국복지패널과 동일하게 저소득층을 과대표집한 대표적인 패널.
- 1968년 시작 이후 약 50년에 걸친 자료 축적.
 - 조사 참여 대상자는 2017년 기준 약 7만 5천 명 이상이고, 최대 6세대 가구까지 포함.
 - 실제 표본에 속한 개인은 2015년 조사 기준으로 2만 4,637명.
 - 총표본은 크게 세 가지로 분류.
 - 경제기획조사(SEO) 표본, 미시간 조사연구센터(SRC) 표본, 이민자 표본이 그것이며 세 가지 표본 각각에 속한 개인은 8,247명(33.47%), 1만 4,151명(57.44%), 2,239명(9.09%).

○ PSID 자료의 활용도

- PSID 자료는 다양한 학문분과(빈곤연구 및 심리학, 의학, 보건학 등)에 걸쳐 널리 활용.
- 예컨대 2017년 기준 PSID 자료를 활용한 연구 실적은 학회지 2,968개에 걸쳐 학술논문 4,412개, 학위논문 802개, 기타 관련 서적 644개로 집계.
- PSID에 대한 주요 후원자는 미국국립과학재단(the National Science Foundation, NSF).

○ PSID의 조사 내용은 시간에 따라 확대

- 1990년대 중반에는 건강, 부, 지출, 자산, 아동 발달, 성인기의 이행, 그리고 시간 사용에 대한 내용까지 포함.
- 조사 영역의 확대로 자료의 활용가능성, 특히 생애주기발달 연구(life course development) 자료로서의 가치가 상승.

제2절 표본 및 가중치⁸⁾

□ 표본

○ PSID의 표본과 그 시간적 변화

- PSID는 소득과 빈곤의 역동성 파악이라는 핵심목표를 달성하기 위해 1968년에 원표본을 구성하였고, 이 원표본은 2개의 하위 표본으로 구성.

8) 표본과 가중치에 대한 내용은 Institute for Social Research(2017)의 PSID User Manuel을 참고하였다.

- 미국 경제기획국(OEO)의 경제기획조사(SEO)에서 추출한 저소득층 1,872가구,⁹⁾ 미시간대학 조사연구센터(SRC)로부터 추출한 비빈곤층 2,930가구.
- 1968년 원표본(PSID gene)은 4,802가구로, 이 가구에 속한 개인은 약 1만 8,000명.

○ PSID 표본은 차수에 따라 증가

- 그 이유는 첫째, 원표본가구에서 출생·입양한 신규 가구원이 표본가구원(sample persons)에 포함되기 때문.
- 둘째, 표본가구의 가구원들이 성장하여 분가하고 경제적으로 독립가구를 구성할 경우 이들도 개별적으로 조사되기 때문.
- 표본가구원의 자녀도 추적조사하는 PSID의 독특한 연구 설계는 표본의 보충, 자료의 전국대표성 확보, 세대 간 결과에 대한 연구 활성화에 기여.

○ PSID의 패널 유지 방식

- 활동패널(Active Panel)에 대한 세 가지 표본 공급의 출처
 - 첫째, 원표본(N), 둘째, 1968년 이후 이민자 대표성을 위해 추가된 1997년 이민자 추가 표본(n), 셋째, 출생, 입양, 결혼으로 원표본에서 발생한 추가 표본(d).
 - 예컨대 1968~2015년 원표본에서 태어난 아동은 매년 약 300~400명으로 추산.
- 원표본에서 분리되어 독립적 가족 단위를 형성하는 성인 자녀

9) 저소득층 과대표집으로, 상당수 빈곤흑인층(African Americans)이 표본에 포함되었다 (Institute for Social Research, 2017).

는 세대 간 요인, 즉 분가(S)에 따른 추가 표본을 구성.

- 패널의 표본 수(P)에 영향을 미치는 그 밖의 흐름들은 크게 증가요인과 감소요인으로 구분.
 - 표본 증가요인은 추적 상태 규정들(f1과 f2)과 그 규정에 따른 추가 신규 가구 발굴 성공(f1과 f2의 흐름의 증가), 표본 손실에 대한 최소화 전략(a1과 a2), 그리고 조사 거부 또는 이전 차수에서 조사되지 못한 가구에 대한 재접촉(r1).
 - 표본 감소요인은 활동패널, 분가, 손실된 표본으로부터의 사망(m1, m2, m3)과 현 활동패널에서의 탈락.
 - 예컨대 1968~2015년 사망으로 인한 표본 탈락은 연평균 약 100명.

○ PSID에서 장기적 표본 숫자의 변화

- PSID의 최초 설계 이후 차수가 지나면서 새로운 출생코호트가 추가되므로 표본 수는 증가.
- 하지만 원표본 자체는 1968년 이후 미국에 온 이민자를 대표하지 못한다는 단점이 존재.
- 이를 극복하기 위해 1968년 이후 미국에 온 이민자표본(post-1968 immigrants) 추가.
 - 1990년 약 2천 가구의 라티노 가구(Latino households) 추가. 이를 통해 일정 정도 이민자 대표성을 확보하게 되었지만 여전히 이민자 대표성을 전체적으로 확보하지는 못함.
- PSID의 표본 규모는 30주년이 되는 1997년에 2배가 됨.
 - 표본 증가에 따른 비용 상승을 억제하기 위해 원표본 중에서 경

제기회조사(SEO) 저소득과대표본(low-income oversample)의 일부가 제외. 이 표본 감소 결정으로 전체 표본 수는 1996년 약 8,500가구에서 1년 뒤인 1997년 약 6,300가구로 축소.

- 동시에 미국 전체 가구 대표성 확보를 위해 1968년 이후 미국에 이민 온 511가구가 이민자표본으로 추가.
- 하지만 예산 제약으로 추가 표본 규모는 제한됨. 결과적으로 이민자 하위그룹 분석을 위한 표본 구성은 충분치 못함.

□ 가중치

○ 종단면 가중치

- PSID의 종단면 가중치는 기본적으로 ‘역(逆)의 확률 가중치(inverse probability weights)’.
- PSID는 초기부터 기존의 여타 조사들과 달리 가구 단위(household-level)가 아닌 가족 기준(family-level)으로 조사가 진행됨.
- 이런 기본 가정과 조사 형식을 바탕으로 가중치가 구성되며, 가중치는 크게 종단면 그리고 횡단면 가중치로 구분.
- PSID의 종단면 가중치는 개인과 가족에 대한 2단계의 작업을 거쳐 완성.
 - 1단계에서 종단면 개인가중치는 조사에 참여한 모든 개인에게 종단가중치를 부여하여 계산.
 - 2단계에서 종단면 가족가중치는 해당 가족에 속한 모든 구성원(원표본에 포함 또는 포함되지 않은 모든 개인)의 개인가중치를 평균하여 계산.

- 1단계에서 해당 개인은 2차례의 연속적 조사(전년도와 금년도 조사)에 대한 참여 성격(여섯 가지 기준)에 따라 1차적으로 10개의 층위로 분류. 2차적으로 그 분류된 층위에 따라 가중치를 부여받음.
- 첫째, 조사에 대한 참여 성격에 따라 해당 개인은 여섯 가지 기준으로 10개의 층위로 구분됨.
 - 6개의 기준은 다음과 같음: 1) 응답(response), 과거 마지막 조사와 현재 조사에 모두 응답한 경우, 2) 최근이탈(recent exit), 과거 마지막 조사와 현재 조사 사이의 기간에 조사에서 이탈한 경우, 3) 과거이탈(past exit), 과거 한 시점에 조사에 참여했으나 과거 마지막 조사 이전에 이탈한 경우, 4) 최근참여(recent entry), 현재 조사에 처음으로 참여한 경우, 5) 재참여/재시도(reentry/recontact), 과거 한 시점의 조사에 참여했지만 과거 마지막 조사에서 이탈했다가 현재 조사에 다시 참여한 경우, 마지막으로 6) 미래참여(future entry), 현재까지 조사에 참여한 적은 없지만 미래의 한 시점에 참여할 경우
 - 6개의 기준 중 응답, 과거이탈, 재참여/재시도, 미래참여의 4개 기준은 각각 1개의 층위를 구성. 나머지 2개의 기준인 최근이탈과 최근참여는 각각 2개(직전 조사와 현 조사 사이에 사망한 경우, 직전 조사에 응답했지만 현 조사에서 무응답한 경우)와 4개(원표본에서 발생한 출생/입양의 경우, 조사이탈한 개인에 의해 출생/입양된 경우, 원표본의 개인과 결혼/동거하는 경우, 그 밖의 이유로 조사에 새롭게 포함되거나 재등장한 경우)의 층위를 구성. 따라서 총 10개의 층위

(=4+2+4)가 구성됨.

- 둘째, 해당 개인은 그가 속한 층위에 따라 가중치 부여.
 - 최근이탈, 과거이탈, 미래참여 집단에 속하는 경우는 현시점 조사에서 무응답한 개인이므로 가중치 0 부여.
 - 조사에 참여했으나 원표본으로 간주되지 않는 개인의 경우도 가중치 0 부여.
 - 재참여/재시도 집단에 속하는 경우는 과거 마지막 조사 시기의 가중치와 동일한 값 부여.
 - 원표본가구에서 출생/입양된 개인은 해당 원표본가구의 가정과 배우자의 가중치 평균값 부여.
 - 기타 표본에 포함된 개인은 해당 특성을 가진 개인에 대한 추출확률의 역수 부여.
- 또한 응답집단(response group)의 대표성 유지를 위해 위에서 계산된 종단면 가중치는 추가적으로 조정.
 - 첫 번째 조정이 필요한 문제는 표본의 손실, 즉 과거 마지막 조사와 현 조사 사이에 생긴 표본 이탈(attrition)과 사망(mortality)에 대한 고려임. 이 표본 손실의 처리를 두고 통상 두 가지 방식, 즉 조정가중치(the adjusted weights)와 이월가중치(the carry-over weights)를 통해 가중치 조정.
 - 전자는 표본 이탈 및 사망으로 인한 표본 손실을 고려한 방식임. 따라서 표본 개인의 가중치는 과거 마지막 조사 가중치에 이탈조정요인(the attrition adjustment factor)을 곱하여 조정.
 - 하지만 후자는 표본 손실을 고려하지 않는 방식임. 따라서

표본 개인의 가중치는 과거의 마지막 조사 가중치가 그대로 적용.

- PSID는 조정가중치 방식을 적용해 옴(예: 1997년 표본 가중치 계산).
- 두 번째 조정이 필요한 문제는 경제적 제약으로 인한 표본 축소 및 이민자 대표성 확보를 위한 표본 추가에 대한 고려임. PSID는 이에 대응하기 위해 다양한 조정 방식을 활용해 옴.
- 예컨대 1997년에 경제기획조사(SEO) 표본의 일부가 제외되어 이를 보정하기 위해 1997년 표본에는 전체적으로 더 큰 가중치가 부여됨. 또한 1997년 이민자표본 추가 과정에서 이민자표본이 예상보다 더 높은 확률로 추출되어 이를 보정하기 위해 추출확률을 반영한 가중치 부여.
- 1997년 핵심표본과 이민자표본을 결합할 때, 전체 표본 내 이민자표본의 비율이 당시 전체 미국 인구 대비 이민자 비율의 추정치(7%)에 부합하도록 표본 가중치 조정.

○ 횡단면 가중치

- PSID의 횡단면 가중치는 위에서 산출한 종단면 가족가중치에 근거하여 2단계를 거쳐 부여.
- 1980년대 초반까지 횡단면 가중치는 별도로 생산되지 않음.
- 1968년 원표본은 아니지만, 나중에 원표본에 속한 개인과의 결혼 등으로 표본으로 진입한 ‘비표본개인(non-sample persons)’의 숫자는 초기에 그리 많지 않았고, 이들의 종단면 가중치는 0으로 부여.
- 가구 특성과 결과에 대한 횡단면 분석에 대해 종단면 가족가

중치 사용이 권고됨.

- 하지만 시간이 지나며 비표본개인의 숫자가 점차 증가하고 표본 내에서 상당한 비중을 차지하면서, 가중치 구성 과정에서 이를 더 이상 무시할 수 없게 됨.
- 1968년 비표본개인의 수는 총표본 개인 1만 7,212명 중 537명(약 3.1%)에 지나지 않았으나 2015년 총표본 개인 2만 4,637명 중 7,132명(약 28.9%)에 달함.
- 따라서 분석을 위한 유효표본 수의 증가를 위해 새로운 개인 수준의 가중치, 즉 횡단면 가중치의 필요성이 제기됨.
- 또한 1984년 미국 소득 및 프로그램참여조사(Survey of Income and Program Participation, SIPP)에서 일하던 통계학자들은 PSID와 같은 동태적/장기적 표본에 진입하는 비표본개인(non-sample persons)을 포함한 가중치 계산 기법을 연구하기 시작.
- 이 흐름 속에서 PSID도 Kalton&Brick(1995)의 연구가 제시한 공정할당방법(fair shares method)으로 불리는 방식을 적용하여 표본개인과 비표본개인을 모두 포함하는 횡단면 가중치를 처음으로 계산.
- 공정할당방법(fair shares method)의 기본가정은 “개별 가구 내 개인을 관찰할 확률은 개별 가구 자체를 관측할 확률과 같다”는 것임.
- 이 가구 균등화와 개별 확률은 1968년 조사된 원표본가구와 개인들에 대해서는 참임.
- 그러나 1968년 이후 진행된 조사에서 1968년 원표본에 속

하지 않았던 비표본개인들에 대한 확률은 알지 못하거나 쉽게 결정하기 어려움.

- 어떤 특정 시점 t 에서 조사된 표본에 대해 PSID 가족 내 개인에 대한 0이 아닌 횡단면 가중치는 공정할당방법에 따라 다음과 같이 부여.
 - 특정 인구집단 추정치(예: 인구총계)의 정확성을 최적화하기 위한 ‘임의적 영향 가중치’를 활용함. 임의적 영향 가중치는 균등한 개인가중치로 정의.
 - 균등한 개인가중치의 기본적 의미는 t 시점에서 표본 내 각 가족은 1968년 원표본이 구성될 때 해당 가족이 선택될 확률과 동일한 것으로 가정하고 가중치 부여.
- 1단계는 기본가중치(base weight) 산출 과정임. 이는 분할표를 이용한 절삭(trimming)과 대체(imputation)를 활용.
 - 기본가중치 산출을 위해 PSID의 표본은 네 가지 기준에 따라 분할표에서 셀로 층화.
 - 네 가지 기준은 다음과 같음: 1) 표본의 소속[조사연구센터(SRC) 표본, 경제기회조사(SEO) 표본, 이민자 표본 중 어디에 속하는지], 2) 가장의 나이(35세 미만, 35~55세 미만, 55세 이상의 3개 연령층 중 어디에 속하는지), 3) 가장의 인종(흑인 여부), 4) 주거지역(북동부, 중서부, 서부, 남부의 4개 지역 중 어디에 속하는지)
- 2단계는 사후층화가중치(post-stratified weight)를 산출하는 과정임. 이는 기본가중치(base weight)에 상시인구조사(CPS, March) 또는 미국지역사회조사(ACS)의 인구학적 통계

를 이용한 조정요인이 고려.

- 사후층화 통제를 위해 1997년부터 2013년까지의 PSID 횡단면 가중치는 상시인구조사(CPS, March Supplement)를 참고하여 조정.
- 2015년 PSID 횡단면 가중치는 미국지역사회조사(ACS)의 공공용 미시자료표본(Public Use Microdatat Sample)을 참고하여 조정.¹⁰⁾

□ 표본추출 방식 및 무응답 처리

○ 표본추출 방식

- 표본추출 과정은 가중치 계산을 위해 면밀한 이해가 필요.
- PSID의 원표본은 두 가지 독립표본, 즉 조사연구센터(SRC) 표본과 경제기회조사(SEO) 표본으로 구성.
- 전자는 미시간대학 조사연구센터(SRC)로부터 층화다단계 추출(Stratified multistaged selection)로 선택된 등확률 표본(equal probability sample)이고, 생산가능인구에 대한 횡단면 표본임. 1968년 2,930가구에 대해 조사 완료.
- 후자는 경제기회조사(SEO) 표본에서 추출한 저소득층에 대한 불균등확률표본(unequal probability sample)임. 1968년 1,872가구에 대해 조사 완료.

○ 무응답 처리

- PSID의 무응답 식별은 다음과 같음.

10) PSID 홈페이지 https://psidonline.isr.umich.edu/data/weights/cross_sec_weights_15.pdf에서 2018. 9. 28. 인출.

- 변수별 유효값의 범위는 다름. 통상적으로 무응답은 '8'=모름(don't know), '9'=조사 거부(refused), '0'=없음(none) 또는 비해당(inapplicable)으로 처리.
- 한편 무응답 처리를 위한 무응답 대체법은 다음과 같음.
 - Duffy(2011)에 따르면 무응답 대체법은 크게 다섯 가지가 있음: 1) 대안변수로의 대체(alternative variable), 2) 과거 직전조사의 자료 이월(prior wave carryover), 3) 집단 평균 대체(group mean replacement), 4) 핫덱 대체(hotdeck replacement), 5) 중위수 대체(median replacement) 방법

제3절 조사 방식 및 관리

□ 조사 방식

- PSID의 조사 방식은 크게 세 가지, 즉 대면조사, 전화조사, 컴퓨터를 이용한 면접조사 방법이 존재.
 - 조사 초기부터 1972년까지는 95% 이상이 면대면 조사로 수행되었고, 이후 전화조사로 진행됨.
 - 1993년부터 컴퓨터를 이용한 전화조사(CATI)로 조사 방법 이행.
 - 2003년부터 설문을 프로그램화하기 위한 블레이즈 소프트웨어(Blaise Software)가 도입됨. 또한 표본과 응답가족에 대한 행정정보 관리를 위해 미시간대학 사회연구소(Institute for

Social Research, ISR)에서 개발한 서베이트락(SurveyTrak)이 사용.

- 2017년 기준 전화조사(phone interview)로 조사 대상자의 85% 이상에 대한 조사가 진행됨. 전화조사가 불가능한 약 15%의 조사 대상자에 대해서는 면대면조사(face-to-face interview)를 진행.

○ PSID의 평균 설문시간은 2015년 기준 90분 정도 소요.

- 90분의 설문시간은 주된 조사 영역에 대한 설문시간(77.7분)과 추가설문시간(13.1분)을 합친 것임.
- 2015년 기준 본조사 영역 내 평균 인터뷰 시간은 고용(employment) 영역에 대한 조사가 18.2분으로 가장 높고, 다음으로 건강 상태, 건강행동, 건강지출(health status, health behaviors, health expenditures) 영역(14.7분), 지출과 소비 영역(11.1분), 소득 영역(8.8분)이 그 뒤를 따름.

○ 설문 내용의 지속적 증가로 인한 조사 대상자의 설문 부담 감소를 위해 2011년부터 다양한 방법을 고안함.

- 예컨대 1968년의 경우 평균 설문시간은 63분이었음. 그러나 2011년의 경우 평균 설문시간은 90분이었음.
- 하나의 방법은 종속형 질문(Dependent Interviewing, DI)을 활용하는 것임. 즉 이전 조사에서 응답한 결과를 가져와 공통된 내용에 대해서는 설문을 대체.
- 또 하나는 설문지 축소(redundant items dropped)임.
- 이러한 노력의 결과, 설문시간은 2011년 90분에서 2015년 78분으로 나타나 12분(13.3%)만큼 줄어듦.

○ 조사의 품질 제고 방법

- 자료의 품질 관리(Quality control check)를 활용.
- 품질관리의 내용은 조사 결과를 재확인하고, 이중 체크, 조사원 체크리스트(Check List) 등의 보조적 자료 활용.
- 자료의 일관성을 위해 PSID 직원들과 학계 연구자들이 정기적으로 자료의 질적 수준 평가.
- PSID 웹사이트의 ‘자료의 질 문헌목록(Data Quality Bibliography)’은 12개의 주제 영역에 대한 자료의 질을 평가한 연구들을 소개.¹¹⁾

○ 사용자 친화적 서비스 제공

- 1996년부터 온라인자료센터(Online Data Center)를 통한 데이터 제공.
- 자료센터를 통해 사용자는 자신의 필요에 맞는 조사 차수 및 변수를 선택하여 분석용 데이터세트를 구성할 수 있음.
- 조사 완료 1~2개월 이후에 원자료(raw data)가 제공됨. 조사 완료 후 약 1년 6개월 뒤에 공식적 자료가 제공됨.

□ 조사 주기 및 조사 응답률

- 조사 주기의 경우 1968년부터 1997년까지는 매년 조사되었고, 1997년 이후 격년으로 조사되고 있음.

11) 소득역동성패널연구(PSID)를 구성하는 12개의 주제영역(topical area)에 대한 자료의 질을 검토한 연구목록을 확인할 수 있다. 12개의 주제영역은 자선, 소비지출, 교육, 보건사 달력, 건강, 가사노동, 소득/빈곤/인구통계변수, 결혼, 연금, 표본추출 및 표본대표성, 시간 사용, 부(wealth) 영역으로 구성된다.(자세한 사항은 다음의 웹주소를 참고할 것 https://psidonline.isr.umich.edu/Guide/Quality/data_quality.aspx)

- 핵심표본의 전체적인 응답률은 1968년 76%에서 2009년 94.3%로 꾸준히 증가함.

□ 비조사기간 패널가구 및 가구원 관리 방식(조사 대가)

- 조사 참여 대상자의 조사에 대한 거부감 완화 및 우호관계 증진을 위해 다양한 방법 활용(McGonagle et al, 2012, p.272).

- 응답자를 위해 다가올 조사의 내용과 시점에 대한 지속적인 공지, 조사의 과학적 가치를 언급하고 조사 참여자의 지속적 참여에 감사를 표하는 정기적(또는 연례적) 뉴스레터 제공, 주소 업데이트를 위한 우편물 발송, 조사 혐오 또는 조사 거부 사례에 효과적으로 대처할 수 있는 특별훈련을 받은 조사원 양성, 인터뷰 직후 현장에서 응답자에 대한 즉각적 인센티브 지급, PSID 자료를 활용한 연구 성과들을 요약한 성과물 발송.

- 표본 마모를 방지하기 위한 다양한 인센티브 제공

- 첫째, 조사 참여 대상자의 지속적이고 일관된 조사 참여를 유도하기 위해 조사답례금 지급.
 - 조사답례금은 면접조사시간 평균 1분당 1달러로 계산되며 이 산식에 따라 지급된 금액은 2018년 기준으로 기존 응답자의 경우 75달러, 신규 응답자의 경우 100달러, 이민자의 경우 150달러. 이 조사답례금은 면접조사 이후 48시간 이내에 우편으로 지급.
- 둘째, 추가적인 정보 제공에 대한 답례로 제공되는 금전적 인센티브가 존재.
 - 해당 응답자가 자발적으로 변경된 자신의 주소지 정보를 조

사탐에 알려줄 경우 답례금으로 10달러 제공.

□ 조직 및 인력: 조사인력, 재정, 관리감독의 세 가지 측면¹²⁾

○ 조사인력 측면

- PSID는 미시간대학 교수진이 총괄적으로 지휘.
- 조사의 체계적인 수행, 관리, 운영을 위해 미시간대학 조사연구센터(SRC)에 조사연구운영팀(Survey Research Operations, SRO unit)을 둬.
- 조사연구운영팀(SRO)은 데이터 수집을 전담하는 조직임. 주된 기능은 조사 설계, 데이터 수집, 데이터 처리에 대한 다양한 서비스를 내부 및 외부 클라이언트들에게 제공.
- 조사연구운영팀(SRO)은 지역 내, 미국 전체, 국가 간 수준에 걸친 다양한 방법론적 조사를 수행함. 특히 국가 간 비교 연구 프로젝트에 대한 상담과 프로젝트 관리 서비스 제공.
- 조사연구운영팀(SRO)은 다중적, 복합적, 연속적인 조사 방식들(modes)에 대한 기반, 능력, 경험을 가짐. 이러한 조사 방식들은 직접(면대면)조사, 전화조사, 우편조사, 온라인조사를 포함함.
- 조사연구운영팀(SRO)의 역할은 크게 아홉 가지로 나누어 살펴볼 수 있음. 일반적인 방법론과 통계적 방법론에 대한 상담, 표본 설계, 가중치와 결측값 대체, 설문지 설계, 프로젝트 관리, 조사원 교육훈련, 데이터의 수집과 처리, 그리고 데이터의

12) 조직 및 인력에 대한 이하의 내용은 PSID 조사연구운영팀(SRO) 웹페이지 <http://www.src.isr.umich.edu/services/survey-research-operations/>에서 2018. 9. 28. 인출.

분석까지 포함함.

○ 재정 측면

- 초기 예산 지원은 미국 상무부 경제기회국(the Office of Economic Opportunity of the United States Department of Commerce)에 의해 이루어짐.
- 현재 연구의 주된 자금지원 출처는 국립과학재단(the National Science Foundation: NSF)임.
 - 2006년 공지된 국립과학재단의 PSID 지원계획에 따르면 2007~2011년 5개년간 국립과학재단의 지원액은 최소 1,200만 달러에서 최대 2,000만 달러(한화 130억~220억 원 상당)가 지출되었을 것으로 추정됨.¹³⁾
- 한편 PSID의 주된 후원자들은 시간에 따라 지속적으로 변화.
 - 지난 수십 년간 PSID의 주된 후원자는 크게 세 주체, 즉 미국 국립과학재단(NSF), 미국 노령연구원(the National Institute on Aging, NIA), 그리고 미국 35대 대통령 존 에프 케네디의 여동생인 유니스 케네디 슈라이버(Eunice Kennedy Shriver)가 설립한 아동건강 및 인간개발연구원(National Institute of Child Health and Human Development, NICHD)임.
 - 가장 최근의 재정 순환주기(financial cycle)인 2012~2016년의 경우 PSID에 대한 재정 지원은 주로 인디애나대학의 릴리 가족 자선학교(the Indiana University

13) NSF 홈페이지 https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=13460에서 2018. 9. 28. 인출.

Family School of Philanthropy), 미국 농무부의 경제연구서비스(the Economic Research Service of the United States Department of Agriculture), 미국 건강인구서비스부의 계획평가 차관보(the Assistant Secretary for Planning and Evaluation of the United States Department of Health and Human Services)로부터 이루어짐.

○ 관리감독 측면

- PSID는 1983년부터 감독위원회(Board of Overseers)의 조언을 받아 옴.
- 감독위원회는 미국 국립과학재단(NSF)이 미국 전체의 학계, 연구자, 정책입안자로 구성된 연구공동체로부터 의견 접수를 활성화할 목적으로 구성.
- 감독위원회 구성원들은 PSID 웹사이트에서 목록 확인 가능.

제4절 조사표의 주요 내용

□ 조사표의 구성¹⁴⁾

○ 일반적 조사 내용

- 1968년부터 고용, 소득, 인구학적 요인에 대한 정보 수집.
- 설문지 내용은 미국 국립과학재단(the National Science

14) 조사표에 대한 이하 내용은 Institute for Social Research(2017)의 PSID User Manuel을 참고하였다.

Foundation, NSF)에 의해 설립된 이사회의 자문 그리고 학계의 수요에 의해 구성되며, 매 시기 새로운 이슈에 대한 학술적, 정책적 도움을 줄 수 있게 구성.

- 2011년 조사의 경우 고용, 수입, 모든 원천별 소득, 전체 가구 소비의 95~100%를 차지하는 지출, 주거, 압류 및 주택 담보대출로 인한 스트레스, 지리공간자료, 건강 상태, 건강 행동, 성인과 그 성인의 아동시기의 건강 상태, 건강보험, 결혼과 출산, 공적 프로그램 참여, 자동차 소유, 부와 연금, 부채 그리고 자선에 대한 정보를 수집.
- 2015년 조사의 경우 인터넷 접근성, 식품안전, 배우자 소득, 건강보험, 부(wealth), 연금 관련 질문 추가.

○ 구체적인 조사 내용

- 주요 설문 내용은 가장 최근의 조사인 2015년 기준 본조사(main interview)를 기준으로 총 열한 가지의 대주제로 구성되며, 15개의 개별 설문지 영역이 존재.
- 열한 가지 대주제는 첫째, 주거, 공과금, 컴퓨터 사용 및 인터넷 접근, 둘째, 고용, 셋째, 가사노동, 식료품비 지출, 식료품지원 프로그램, 교통비, 교육비, 기타 지출, 넷째, 소득, 다섯째, 건강 상태, 건강행동, 건강 관련 지출, 여섯째, 결혼과 출산, 일곱째, 새로운 가구주와 배우자/파트너의 배경, 여덟째, 자선적 기부와 자원봉사, 종교, 도움받은 것, 아홉째, 연금, 열째, 재정적으로 어려운 시기의 소득 및 공적부조, 열한째, 부와 적극적 저축임.

○ 조사표의 주요 내용을 가구 일반 사항, 소득, 지출, 자산, 부채, 복

지, 노동, 보건, 기타 영역의 다섯 가지를 중심으로 하여 살펴보면 다음과 같음.

- 첫째, 가구 일반 사항의 측면에서 PSID 표본과 관련된 용어와 정의는 1968년 미국 통계청(the Census Bureau)으로부터 채택.
 - 1968년부터 사용해 온 정의와 용어들은 일관성(consistency)과 직관적 추적규칙(straightforward following rules)을 유지하기 위해 변동 없이 거의 그대로 사용.
 - 응답자 성인 개인은 단독 1차 성인(a single primary adult)이라는 용어로 표현함. 혼인한 가족(a married-couple family) 내에서 가구주(the Head)는 일반적으로 남편(the husband)으로 정의.
 - 응답자들은 그들 자신과 그들의 배우자/파트너(spouse/partner), 그리고 함께 살고 있는 모든 다른 가족 구성원으로 구분되며, 이 모두는 가족 단위(Family Unit, FU)로 지칭.
 - 기타 가족 단위 구성원(Other Family Unit Member, OFUM)은 해당 조사 시점에서 가구주 혹은 배우자가 아닌 응답자의 가족 구성원이고 응답한 가족과 함께 살거나 시설(예컨대 대학교, 감옥, 병원, 혹은 군대)에 일시적으로 거주하는 경우를 포함.
 - PSID는 조사의 시계열적 일관성을 극대화하기 위해 차수별로 동일한 가족 구성원을 면접하려고 지속적으로 노력한 결과, 연속된 차수 간에 약 95%의 가족들이 그 두 조사에서 동일한 응답자를 가짐.

- 또한 PSID는 표본 대표성 확보를 위해 표본 구성원들의 가구 변동 사항이 있을 경우 표본 구성원들을 추적함. 추적 정보에는 표본 구성원들, 그리고 그들과 동일한 가구에 거주하고 있는 다른 가족 구성원들에 대한 정보가 포함됨.
- PSID 가족 단위(FU)에서 나간 가족 구성원은 만약 그들이 표본 구성원이면서 다른 독립된 가구에 거주할 경우 분리된 가족 단위(separate family unit)로서 면접할 수 있는 자격이 있음.¹⁵⁾
- 둘째, 소득 및 지출의 측면에서 ‘총가족소득(Total family income)’은 조사 전년도 기준 세전소득 개념으로 계산되고, PSID 가족 전체 구성원들의 다양한 소득원천들의 합으로 구성.
 - 가장 최근의 2015년 가족 파일(Family File)을 기준으로 총가족소득은 다음의 일곱 가지 소득요소들의 합.
 - 2015년 가족 파일(Family File)에는 이 합계액이 총가족소득(Total Family Income-2014)이라는 이름에 변수명 ER65349로 저장.
 - 이 총가족소득은 과세 대상 연도 2014년 기준으로 조사된 것이며 부(-)의 소득을 포함. 부(-)의 소득은 사업 또는 농업 손실의 결과로 발생한 순손실(net loss)을 의미. 부(-)의 소득은 1994년 이전 차수들에서는 1달러로 하한 코딩(bottom-coding)됨.
 - 총가족소득은 세전소득이기 때문에 세후소득을 구하기 위

15) 2015년 소득역동성패널연구(PSID)에 대한 자료 압축 파일을 다운받은 후, 코드북("FAM2015ER_codebook.pdf")을 참조.

해서는 세금 및 사회보장분담금에 대한 계산이 필요.

- PSID 자료로부터 소득 및 사회보장세(income and payroll taxes)를 계산하기 위해 택심(TAXSIM)이라는 프로그램 사용.¹⁶⁾
- 소비지출의 측면에서는 1999년 소비지출을 포함한 이후 차수마다 조사함. 조사 내용은 열한 가지 세부 영역, 즉 식료품비, 주택 수리 및 유지비, 가구집기 및 가사용품비, 피복신발비, 여행 및 휴가비, 교양오락비, 의료비, 주거비, 교통비, 교육비, 아동양육비로 구성.¹⁷⁾
- 셋째, 자산과 부채의 측면에서 살펴보면 PSID는 1984년 부(또는 재산) 모듈(wealth module)을 포함하였고, 이후 계속 조사.
 - 지난 2년간 가족구성원 중 상당액의 선물(gifts) 또는 유산 상속(inheritance)을 받은 적이 있는지, 만일 받았다면 받은 것 중 가치가 큰 순서대로 세 가지를 조사.
 - 2013년에는 상당액의 선물 또는 유산상속을 특정 연도에 받았는지만 조사하였고, 2015년에는 각 세부 유산상속이 몇년도에 일어났는지도 조사.
 - 순재산(net worth)은 설문지의 주거 영역과 부(또는 재산) 영역의 정보를 통해 계산. 단기간 부채(short-term debts)를 포함한 아홉 가지 유형의 자산/부채를 조사.

16) NBER 홈페이지 <http://users.nber.org/~taxsim/to-taxsim/psid-kks/>에서 2018. 9. 28. 인출.

17) PSID 홈페이지 <https://psidonline.isr.umich.edu/data/sl/Consumption-PSID.pdf>에서 2018. 9. 28. 인출.

- 넷째, 복지 및 노동 측면에서 고용 영역(설문지 BCDE)을 통해 가구주와 배우자의 직업에 대한 세부 정보 조사.
 - 조사 내용은 노동한 주 수, 상병(illness) 또는 실업(unemployment)으로 노동하지 않은 주 수, 총노동시간과 소득, 가구주와 배우자의 주된 직업, 세 가지 추가적인 부업을 포함.
 - 다섯째, 보건 및 기타 영역의 측면에서 건강 상태 및 행동의 열한 가지 영역을 조사.¹⁸⁾
 - 일반적 건강 상태, 일상생활행동/도구적 일상생활행동(ADL/IADL-Measures), 구체적 건강 상태, 아동기(17세 이전) 건강 상태, 정서적 복지와 정신건강, 노동/활동 제약들, 일하지 못한 기간 및 요인, 건강행동, 영양시설 이용 경험, 건강 관련 공적소득이전 프로그램 참여, 부모의 건강 상태를 조사함. 마지막으로 기타 영역에는 식품, 영양, 체질량 지수(food, nutrition, BMI), 주거(housing), 교육(education) 등이 있음.
- 생애과정연구(life course research)를 위한 2개의 모듈 추가.
- 첫째, 아동기 건강달력(childhood health calendar)은 2007년에 추가.
 - 아동기 건강달력 모듈은 건강 상태와 사회경제적 지위 사이에 강하고 지속적인 상관관계에 대한 원인들, 그리고 생애과정에 걸친 그 상관관계의 변화와 관련한 자료를 제공할 목적

18) PSID 홈페이지 <https://psidonline.isr.umich.edu/data/sl/Health-PSID.pdf>에서 2018. 9. 28. 인출.

으로 설계.

- 이 모듈은 가구주와 배우자가 그들의 아동기 시절 다음과 같은 건강 상태들(예컨대 천식, 비만, 호흡기 장애, 언어 장애, 우울증, 약물 또는 알코올 문제, 그리고 다른 정서적 문제들)을 경험했는지, 경험했다면 그 시기가 언제였는지에 대한 회상적 설문 포함.
 - 이러한 자료의 수집은 생애주기의 매우 초기 단계에서 나타났던 지표들(markers)을 제공함으로써 건강이 생애과정발달에 미치는 효과를 검토하기 위한 PSID의 연구 역량을 증진.
 - 둘째, 주택담보대출로 인한 스트레스(mortgage distress)는 2009년에 새롭게 추가.
 - 주택담보대출로 인한 스트레스 모듈은 압류활동, 지불체납, 주택담보대출 변경, 이듬해 주택담보대출 상황 어려움에 대한 예상 등 관련한 새로운 질문들을 포함.
 - 이러한 자료는 최근의 경기 침체가 가족 형성과 해체, 출산과 교육 성취를 포함하는 일련의 결과들에 미치는 중기적 효과(medium-term effects)를 검토하기 위해 사용.
 - 예컨대 연구자들은 더 높은 실업률, 소득과 부의 감소가 청년층(young adults)으로 하여금 그들의 고등교육 성취, 결혼과 동거, 그리고 출산을 연기하거나 기피하게 하는지에 대해 조사 가능.
- PSID는 조사 영역을 확대해 왔을 뿐 아니라 특정 주제에 대한 부가조사도 진행.
- 이 부가조사는 크게 다섯 가지로 나뉨.

- 첫째, 장애와 시간사용 보충조사(Disability and Use of Time Supplement, DUST), 둘째, 아동발달 보충조사 및 성인기로의 이행 보충조사(The Child Development Supplement and the Transition into Adulthood Supplement, CDS and TAS), 셋째, 웹 보충조사(Web Supplements)로서 아동기 회상적 상황연구(Childhood Retrospective Circumstances Study, CRCS), 넷째, 웹 보충조사로서 복지와 일상생활 보충조사(Wellbeing and Daily Life Supplement, WDLS), 다섯째, 세대 간 이전(Intergenerational Transfers) 보충조사가 그것임.

제5절 패널 활용

□ 패널을 활용한 연구 실적

○ 네 가지 지표로 살펴본 자료 활용 실적

- 첫째, PSID 자료를 이용한 출판 실적
 - 2017년 1월 기준 4,412개의 전문적 심사를 받은 출판물(peer-reviewed publications)이 발간된 것으로 집계.
 - 이 4,412개의 출판물은 2,966개의 학술논문, 644권의 책과 책의 장, 802개의 박사학위 논문을 포함.
- 둘째, PSID 자료 활용 연구에 대한 미국 국립보건원(NIH)과 국립과학재단(NSF)의 연구비 지원금(grants) 규모
 - 2017년 4월 기준 미국 국립보건원(NIH)과 국립과학재단(NSF)은 PSID 자료의 수집 및 이차적 자료분석(secondary

data analysis)을 지원할 목적으로 약 300개의 연구지원금(awards)을 조성.

- 셋째, 등록된 사용자 수, 웹사이트 활동(방문 횟수), 자료 다운로드 횟수
 - 2015년 기준 11만 3,596명의 고유 방문자들(이는 IP주소로 확인)이 약 430만 건의 웹사이트 방문 횟수 기록.
 - 자료 다운로드는 두 가지 방식, 즉 사용자가 직접 자료센터에서 맞춤형 데이터세트(customized dataset)를 만들거나 압축된 파일(zip file) 형태로 전체 파일을 다운로드하는 방식 두 가지가 있음. 2015년 기준 이 두 방식에 따른 총다운로드 건수는 3만 6,566건임.
 - 2015년 기준 2만 5,267명의 사용자가 등록됨. 2014년과 비교할 때 2015년 등록자 수는 10%(즉 2,500명 이상의 사용자) 증가함. 총 등록 사용자의 69%는 경제학 전공자, 8%는 사회학 전공자, 그리고 나머지 13%는 기타 전공자(교육학, 심리학, 인구통계학, 아동발달학, 의학, 지리학, 그리고 기타 전공)임.
- 넷째, 사용자 학술대회 개최
 - 사용자 연례 학술대회(PSID Annual User Conference)를 통해 PSID 자료를 활용한 연구논문을 공모하고 그 발표를 위한 장을 마련.
 - 학술대회의 주목적은 PSID 자료를 사용한 연구자들이 연구 성과를 발표하고 경륜이 있는 연구자들과 직원들로부터 환류(feedback)를 받기 위함.

- 학술대회로 조사 영역의 양적 확대 및 질적 성숙을 꾀하는 동시에 국립보건원(NIH) 및 국립과학재단(NSF) 등 후원기관들로부터 미래 재정 지원에 대한 근거자료를 마련할 수 있는 기반을 조성하려 노력.

제6절 복지패널에 주는 시사점

□ PSID 자료 검토의 의의

- PSID 자료 검토의 주목적은 한국복지패널 자료의 조사 내용, 조사 방식, 조사 관리에 대한 시사점을 얻고 향후 개선 방향을 모색하기 위함임.

○ PSID 자료 검토가 한국복지패널에 주는 시사점

- 첫째, 세계화와 그로 인한 빈곤과 불평등의 증가, 그리고 그에 대응하는 복지국가별 대응전략이라는 커다란 연구적 관점에서 패널의 생존전략 중 하나의 요점은 바로 ‘패널의 국제화’에 있고 PSID는 이를 위해 CNEF 프로젝트를 수행하고 있다는 점임.
- 한국복지패널도 PSID와 마찬가지로 국가 간 비교 연구를 위한 자료구축 네트워크에 적극 참여할 필요가 있음.
- CNEF는 PSID를 포함한 8개국(영국 BHPS, 캐나다, 호주 HILDA, 독일 SOEP, 네덜란드, 스위스, 한국 KLIPS, 러시아)의 패널자료와 연계됨. 최근에는 일본 게이오대학과 연계하여 일본의 패널자료도 포함될 예정임.

- ‘패널의 국제화’의 관점에서 한국복지패널이 CNEF 프로젝트에 참여하기 위한 절차는 세 가지 단계로 요약됨.¹⁹⁾ 첫째, CNEF 구축팀에 영문으로 된 자료 설명(변수, 샘플링, 조사 절차, 표본 수, 표본탈락률, 그리고 관련된 설명자료) 제출, 둘째, 한국복지패널이 CNEF에 균등화된 자료(harmonized data)로 포함될 경우 연구자들의 수요가 어느 정도이고 연구자들이 얻는 이득이 무엇인지에 대한 설명 제출, 셋째, 위의 두 단계를 토대로 CNEF 구축팀은 한국복지패널을 검토하고 이를 통해 국가 간 비교가 가능한 균등화(harmonized) 변수들을 생성할 수 있는지를 검토.
- 둘째, 사회구조적 요인의 변화와 그에 따른 복지욕구의 변화에 대한 반응성(responsiveness) 제고 차원에서 ‘패널의 조사 내용 다각화 및 타 자료와의 연계성 강화’가 필요.
- PSID는 2007~2010년 경제위기가 미국에 미친 영향, 특히 주거와 관련된 영향을 파악하기 위해 시기적절하게 주택담보대출 및 그로 인한 스트레스를 부가조사 형식으로 조사.
- 이러한 전통적인 경제적 위험과 더불어, 인간 자체의 생애주기적 위험, 특히 생애주기 초기의 결정론적 위험에 대한 인식의 중요성에 착안하여 아동기 성장환경과 관련된 요인과 유전적 요인(예컨대 타액 시료와의 결합 자료) 또한 부가조사 형식으로 조사.
- 이러한 PSID의 시의적절한 부가조사는 한국복지패널에 유의미한 시사점을 줌.

19) 이하의 내용은 2018. 4. 13. CNEF 자료구축 프로젝트의 연구책임자인 Dr. Lillard에게 문의한 결과를 참고하였다.

- 예컨대 최근 미세먼지 등의 환경적 위험과 이 위험이 빈곤지역에 집중되는 경향이 강해지면서 환경-빈곤-건강 불평등에 대한 인과적 관계가 새로운 이슈로 부각.
- 이러한 관점에서 한국복지패널의 조사표도 환경-빈곤-건강의 인과적 결합을 밝힐 수 있는 방향으로 조사 영역을 확대하거나 타 자료와의 연계성 강화가 필요.
- 물론 패널조사의 일관성과 지속성, 그리고 예산 제약과 시간 제약 속에서의 자료 생산이라는 압박을 생각한다면 현재 지속되는 핵심조사에 새로운 이슈에 대한 조사 내용을 적극적으로 반영하기는 힘들 것임. 하지만 패널 구축이 사회 변화에 대한 능동적 대처 및 정책 대안 모색을 그 실천적 목표로 한다는 점을 생각해 볼 때 한국복지패널은 조사 내용의 다각화 및 타 자료와의 연계성 강화를 좀 더 적극적으로 검토할 필요가 있음.

제 4 장

독일 Socio Economic Panel(SOEP)

제1절 조사 배경

제2절 표본 및 가중치

제3절 조사 방식 및 관리

제4절 조사표의 주요 내용

제5절 패널 활용

제6절 복지패널에 주는 시사점

4

독일 Socio Economic << Panel(SOEP)

제1절 조사 배경²⁰⁾

□ 독일 리서치연합센터(German Research Foundation's Collaborative Research Center)는 1983년 센터의 사업으로 '사회정책의 근거를 마련하기 위한 미시적 수준의 데이터 확보' 프로젝트를 기획함. 이는 SOEP 패널이 독일의 공식 가구패널로 정착되는 과정의 첫 번째 활동이 됨. SOEP 패널 도입의 연혁을 개관하면 다음과 같음.

○ 1983년 센터 사업에서 프랑크푸르트대학과 만하임대학이 연합하여 가구 패널조사의 기초를 설계함. 당시 프로젝트 리더는 크루프 교수(Prof. Hans-Juergen Krupp)로 1988년까지 독일 가구 패널조사를 총괄함.

○ 이후 2010년까지 바그너 교수가 프로젝트 리더로 조사를 총괄함. 2011년 이후 독일 가구 패널조사를 DIW Berlin이 매해 실시하는 것으로 확대, 변경되면서 이전의 프로젝트 리더는 조사를 총괄하는 DIW Berlin의 행정 이사직으로 합류하게 됨. 그리고 베를린대학의 슈프 교수(Prof. Juergen Schupp)가 SOEP 패널조사의 총책임을 맡으면서 SOEP 패널조사에 베를린대학이 합류하게 됨.

○ 현재 크루프 교수, 바그너 교수, 슈프 교수는 SOEP 패널조사에

20) 이하 내용은 강석훈(1997, pp.5-11, 재인용)을 참고하였음.

관한 다양한 보고서를 집필하면서 SOEP 품질 관리에 기여하고 있음.

□ 독일은 유럽에서 최초로 1984년 가구패널조사인 SOEP를 시작함.

○ SOEP는 독일 인구 및 소득 통계의 수립을 목적으로 함. 고용, 교육, 자산, 소득, 개인이 가지고 있는 정부 정책에 대한 가치관 등을 매해 조사하고 있음(강석훈, 1997, p.5).

○ SOEP의 목적은 개인, 가구의 객관적인 생활환경, 주관적인 삶의 질, 삶의 변화 등과 관련된 다양한 측면의 자료를 제공하는 것임. 이는 궁극적으로 독일 사회 전반을 통계를 활용하여 다목적으로 분석할 수 있도록 자료를 제공하는 것에 목적을 둬.

○ 이를 통해 SOEP 데이터 분석 연구는 정부 정책의 기본 방향을 결정하는 중요한 근거자료로 활용되고 있음.

○ 또한 SOEP에는 국가 간 비교를 위해 기준점 등의 장치가 설계되어 있음.

□ SOEP는 1990년 6월부터 옛 동독지역의 가구를 표본으로 추가하여 조사하고 있음.

○ 1990년까지 서독 지역에 국한하여 패널조사를 하였으나 1990년 6월부터 동독 가구를 표본에 추가하여 조사를 실시함.

- 통일에 대비하기 위해 통일 6개월 이전부터 동독 가구를 표본으로 추가하여 조사를 실시하였음(강석훈, 1997, p.5).

○ 또한 통일 이후 많은 이민자가 발생하여 1994년과 1995년에는 이민자 표본을 추가하였음.

- 특정한 주제가 조사될 필요성이 있을 경우 특수한 샘플(Innovation Sample: SOEP-IS)을 추가하는 방식.

- SOEP를 주관하는 공식 기관은 독일 주요 경제연구기관인 DIW Berlin(Deutsch Institut für Wirtschaftsforschung Berlin)임.
 - 현재 DIW Berlin의 주관하에 프랑크푸르트대학, 만하임대학, 베를린대학이 공동으로 조사연구를 진행하고 있음.
 - 실제로 현장 면대면 조사는 학술조사 전문의 민간기관인 TNS 인프라테스트(TNS Infratest Sozialforschung)에서 수행하고 있음(강석훈, 1997, p.6).
- 독일 정부의 재정지원으로 수행되는 SOEP 패널은 유럽 각국에서 실시하는 가구 패널조사의 모태가 되었다는 평가를 받고 있음.
 - SOEP는 미국의 PSID를 기반으로 시작되었으나 설문지 구성 방식이나 조사 방식에 관해서는 PSID보다 선진화된 패널조사라는 평가를 받고 있음.
 - SOEP는 초창기 DIW의 예산이 지원되었지만 1990년 종료되었고 이후 1995년까지 국가예산위원회의 지원을 직접 받음(강석훈, 1997, p.5).
 - 현재는 독일 과학재단(German Science Foundation)의 재정 지원으로 조사가 진행되고 있으며, 주관기관인 DIW는 연구 시설과 인력을 지원하고 있음.

제2절 표본 및 가중치

1. 표본²¹⁾

- SOEP는 조사 연도와 조사 대상 등이 매우 다양하여 표본의 구성이 복잡한 형태를 이루고 있음.
 - SOEP의 핵심 표본(core study)을 중심으로 여러 종류의 SOEP-IS (Innovation sample) 등이 있음.
 - 다른 조사는 SOEP 핵심 표본을 다시 랜덤 샘플링한 조사라고 할 수 있어 본 연구에서는 SOEP 핵심 표본을 중심으로 표본 및 가중치를 살펴보고자 함.
- 1차 연도 조사 당시 1984년 표본 A, B로 시작해 1990년도에 표본 C를 추가함.
 - 표본 A는 통일 전 서독 지역 독일인 표본 4,528가구임.
 - 표본 B는 스페인, 이탈리아 등 외국 국적을 지닌 가구주로 구성된 표본이며 규모는 1,395가구임.
 - 표본 C는 독일인이면서 동독에 거주하고 있는 사람을 표본으로 함.
- 표본 A를 추출하는 데 사용된 추출 방식은 1982년 독일 경영조사연구소(German Marketing Research Institute)의 연구팀 ADM (Arbeitsgemeinschaft Deutscher Marktforschungsinstitute)이 개발한 ADM 마스터 테이프(ADM-master tape)를 바탕으로 함.

21) SOEP의 표본은 이상호(2006)와 Anne B, et al(2018)을 참고하였음.

- 즉 다단계층화추출 방법을 활용하여 584개의 표본 포인트를 선정함. 이러한 표본은 층화 단계의 첫 번째 단계를 의미하며, 우리나라 통계청에서 사용하는 조사구(Enumeration District: ED)와 비슷한 개념임.
 - 표본이 추출되면 면접원들이 주소를 작성하고, 계통추출 방법을 활용하여 약 12개의 표본가구를 조사구에서 선정함.
- 표본 B는 스페인, 이탈리아 등 외국 국적을 지닌 가구주로 구성된 표본이며 할당임의추출 방식으로 추출함.
- 즉 표본을 추출할 지역을 선정하고 해당 지역으로 이민한 사람이나 거주하고 있는 사람의 정보를 이용해 할당임의추출을 수행함 (강석훈, 1997, p.9).
- 표본 C는 독일인으로 구성된 표본이며 동독에 거주하는 독일인으로 구성됨. 이는 서독에 거주하는 독일인 표본과 비교하기 위해 표본이 확장된 것임.
- 표본 추출에는 동독의 중앙거주자목록(central resident's file)을 사용함.
 - 동독 지역의 크기에 따라 마스터(master)표본을 층화하고, 360개의 표본 포인트를 작성한 후 표본 포인트별로 가구 수에 비례하여 표본 수를 할당하고 계통추출 방법을 실시함.
- 표본 D는 1994·1995년도에 시작된 이민자 표본으로 표집 시기에 따라 D1(236가구)과 D2(295가구)로 구성됨.
- 1994·1995년도 표본 B에 포함되지 않은 이민자를 포함하기 위

해 추가한 표본.

- 가구원 중 외국에서 서독으로 거주지를 옮긴 가구원이 있는 가구로 구성하며 기간은 1984년 이후로 한정함.
- 표본 E는 그동안 이민, 사망 등으로 탈락한 패널 가구를 보완하기 위해 1998년 독일 거주민 전체를 모집단으로 하여 추가한 것임.
- 기존의 표본들과는 독립적으로 표본 A와 유사한 방식으로 추출함.
- 표본 F는 독일 정책의 효과성을 평가하기 위해 무자녀 가구, 복지수급자 가구 등을 모집단으로 하여 추가한 것임.
- 2000년 추가되었으며, 독일인과 이주자 모두를 포함하는 표본으로 추가되었음.
- 표본 G는 고소득자 표본이 과소표집되었다는 판단하에 소득 상위 2.5% 이내의 가구를 추가한 것임.
- 본래 소득 양극단 표본이 상대적으로 적다는 판단하에 고소득자와 극빈계층 양 집단의 표본을 모두 추가하고자 하였으나 거주지가 일정하지 않은 빈곤층 등은 추적조사가 어렵고 조사 결과에 대한 신뢰성을 확보하기 어렵기 때문에 소득이 높은 표본에 대해서만 추가적으로 표집을 실시함.
- 2002년부터 월소득이 7,500DM(3,835유로)보다 높은 가구를 대상으로 표집을 실시하고 있음. 2009년부터는 소득 기준을 조정하여 4,500유로보다 높은 가구를 대상으로 추적조사를 실시함.

- 2000년대 후반부터는 그동안의 패널 탈락자를 보완하기 위해 2006년 표본 H, 2011년 표본 J, 2012년 표본 K가 추가됨.
- 2013년에는 이민자 표본을 보완하기 위해 표본 M1(2,723가구)을 추가함.
- 표본 L1은 코호트 표본(cohort sample)으로 이전에 조사된 ‘독일 가족조사(Familien in Deutschland: FiD)’를 다시 시행하는 형태.
- 이 조사는 가족, 청소년, 여성 관련 연방 부처의 정책효과성 분석을 목적으로 함.
- 가족의 아동을 연령 코호트로 구분하여 조사하고 있음.
- 2010년 1,247가구를 시작으로 SOEP 조사 표본에 통합됨.
- 표본 L2/3는 스크린 표본(screening sample)으로 이것 또한 이전에 조사되었던 ‘독일 가족조사(Familien in Deutschland: FiD)’를 다시 시행하는 형태.
- 이 조사는 독일 가족정책의 각 영역, 즉 일인가구, 한부모가구, 다자녀가구, 저소득가구에 관한 것임.

〈표 4-1〉 독일 SOEP의 표본 특성 개관

표본	시작 연도	가구	개인	표본
A, B	1984	5,921	16,205	서독 거주자(A: 독일인 가구, B: 외국인 가구)
C	1990	2,179	6,131	동독에 거주하는 독일인 가구
D1	1994	236	733	1984년 이후 서독으로 이주한 가구
D1/D2	1995	522	1,665	
E	1998	1,067	2,470	서독에 거주하는 독일인 가구
F	2000	6,052	14,525	독일에 거주하는 독일인 및 이주민 가구
G	2002	1,224	3,538	월 소득이 7,500DM(3,835EURO) 이상인 가구
H	2006	1,506	2,616	독일에 거주하는 독일인 가구
J	2011	2,616	5,161	
K	2012	1,526		
L1	2010	1,247	2,311	독일 가족정책 평가 조사 가구 (아동 연령별 코호트 조사)

주: L2/3 표본은 스크린 표본으로 조사 설계가 상이해 여기에서 다루지 않음(CATI-CAWI-CAPI 모델).
자료: 이상호(2006), p.68; Anne B, et al(2018), pp.2-4.

□ 이러한 지속적인 표본 추가의 결과 SOEP로 현재 1만 가구 이상의 표본이 관리되고 있음. 2016년 조사 결과를 통해 전체 표본수와 가구, 성인, 청소년, 아동 집단별 표본 수를 개관하면 다음의 〈표 4-2〉와 같음.

〈표 4-2〉 표본 집단별 가구, 가구 구성원 조사 대상의 수

(단위: 가구, 명)

표본 집단	가구 수	성인의 수	청소년 수	아동 수	총 응답자 수
A 1984~	1,629	2,719	30	84	2,833
B 1984~	228	429	9	13	451
C 1990~	1,073	1,726	24	53	1,803
D 1995~	173	288	4	7	299
E 1998~	68	102	-	1	103
F 2000~	2,094	3,421	34	56	3,511
G 2002~	590	1,037	6	9	1,052
H 2006~	639	1,052	16	23	1,091
J 2011~	1,883	3,058	38	68	3,164
K 2012~	1,046	1,675	24	31	1,730
L1 2014~	1,122	2,055	36	153	2,244
전체	10,545	17,562	221	498	18,281

자료: Anne B. et al(2018), p.5.

○ 2016년 SOEP 표본 A부터 L1까지의 표본은 총 1만 545가구가
 며, 여기에 총 1만 8,281명의 가구 구성원이 개인 표본으로 있음.
 총가구원 중 1만 7,562명이 성인이며 221명이 청소년, 아동(초
 기 청소년), 즉 대학 이하의 초, 중, 고등 과정 재학생이 498명임.

□ 한편 SOEP가 도입된 1984년부터 가장 최근 조사인 2016년까지
 응답자 규모의 변화를 살펴보면 다음의 〈표 4-3〉과 같음.

〈표 4-3〉 1984~2016년 표본 집단 규모(명)

응답자 수	1984	1990	1995	1998	2000	2002	2006	2011	2012	2014	2015	2016
A+B	12,239	9,518	8,798	8,145	7,623	7,175	6,203	4,541	4,204	3,761	3,497	1,857
C	-	4,453	3,892	3,730	3,687	3,466	3,165	2,392	2,262	2,006	1,853	1,726
D	-	-	1,078	885	837	780	684	461	435	365	337	288
E	-	-	-	1,932	1,549	1,373	1,199	961	160	128	110	102
F	-	-	-	-	10,886	8,427	6,997	4,984	4,610	4,049	3,773	3,421
G	-	-	-	-	-	2,222	1,801	1,358	1,285	1,168	1,089	1,037
H	-	-	-	-	-	-	2,616	1,478	1,392	1,259	1,162	1,052
J	-	-	-	-	-	-	-	5,161	4,229	3,498	3,279	3,058
K	-	-	-	-	-	-	-	-	2,473	1,962	1,815	1,675
L1	-	-	-	-	-	-	-	-	-	2,311	2,211	2,055
전체	12,239	13,971	13,768	14,692	24,582	23,443	22,665	21,336	21,050	20,507	19,162	17,562

자료: Anne B, et al(2018), p.7.

○ 표본의 규모를 보면 1984년부터 2000년까지 1만 2,239명에서 2만 4,582명으로 증가하다가 2002년부터 점차 감소하여 2014년에 1만 9,162명으로 그리고 가장 최근 조사인 2016년에는 1만 7,562명으로 감소하였음.

□ SOEP 패널의 안정성(Panelsstabilitaet)을 2008년부터 2016년까지의 표본 집단별로 살펴보면 다음의 [그림 4-1]과 같음.

[그림 4-1] 2008~2016년 SOEP 표본 집단별 안정성



자료: Anne B, et al(2018), p.9.

- 패널의 안정성은 지난해 대비 표본의 규모를 의미함. SOEP 패널의 안정성은 적어도 5년마다 관리하여 이탈 가구를 보완하도록 관리하고 있음.
- 2016년 패널의 안정성은 A~H 표본의 경우 평균 93%로 나타남. 이 중 B 표본(서독 지역 외국인 대상)의 안정성이 85.3%로 가장 낮았고, G 표본(고소득자 대상)의 안정성이 97.4%로 가장 높았음.
- 상대적으로 최근의 표집 집단인 J와 K와 L1의 안정성은 각각 95%, 94.4%, 94.8%로 표본이 생성된 후 2016년 평균 이상의 안정성을 보임.

2. 가중치²²⁾

- 조사 첫해에 참여한 대상을 원표본 가구원(original sample member: OSM)이라고 하며, 두 번째 조사 때부터 결혼이나 출생, 기타 전입 등으로 구성원이 증가할 경우 비표본 가구원(nonsample member)으로 분류함(백지선·심규호, 2012, p.3).
- 현장에서 비표본 가구원과 접촉했을 때 어떻게 대처해야 하는지에 대한 규정을 세울 필요가 있는데, 이것을 추적규칙(follow-up rule)이라고 함(백지선·심규호, 2012, p.3).
- 추적규칙에 따라 패널에서 나가거나 패널에 들어오는 경우가 발생하게 됨. 원표본 가구원과 추적조사할 비표본 가구원을 포함한 확장된 원표본 가구원을 영구표본 가구원(permanent sample member: 이하 PSM)이라고 함. 영구표본은 패널조사에 따라 정의하는 바에 차이가 있음(백지선·심규호, 2012, p.3).
- SOEP에서는 1차 연도에 응답한 OSM, PSM의 출생/입양 자녀, 최근 이민자, PSM 자녀가 있는 파트너, PSM 자녀가 없는 파트너, 기타 가구원 모두를 PSM으로 간주함.
- 즉 SOEP는 패널 이탈로 인한 표본 수 감소를 보완하기 위한 추적규칙으로 조사가 거듭되면서 접촉하게 되는 모든 대상자를 포함하는 접촉모형(contagion model)을 적용하고 있음.
- SOEP의 1차 웨이브 가구가중치와 가구원가중치는 다음과 같음(백지선·심규호, 2012, p.16).

22) 백지선·심규호(2013)의 독일 SOEP 패널 관련 내용을 요약한 것임.

○ 가구의 설계가중치에 무응답 조정과 벤치마킹 조정을 실시하여
가구가중치 산출.

○ 가구가중치를 모든 가구원가중치로 부여.

□ SOEP의 2차 웨이브 이후 가구가중치와 가구원가중치는 다음과 같음.

○ 가구가중치

- 초기 가구가중치로 1차 웨이브 최종 가중치를 부여.
- 신규 가구원을 반영하기 위해 초기 가구가중치를 조정한 모형을 적용.
- 조정된 가구가중치에 무응답 조정과 벤치마킹 조정을 실시.

○ 가구원가중치

- 최종 가구가중치를 모든 가구원가중치로 부여.
- 무응답 조정과 벤치마킹 조정을 실시.

□ SOEP는 횡단가중치 조정 방법으로 모형 방법(가구가중치 조정 →
가구원가중치)을 적용하고 있음. 이와 관련하여 가구 변화 요인과
가중치 처리 방법은 다음과 같음(백지선·심규호, 2012, p.18).

○ 일반 진입 가구원(regular household entrants)

- 처음 패널조사가 시작될 때 모집단 안에 있었으나 표본으로 선정되지 않았다가 추가 표본으로 진입한 가구원. 출생이나 입양, 이민자 등과는 구별됨.
- 가구가중치와 가구원가중치를 감소시키는 영향을 미침.

○ 출생(birth) 및 입양(adoption)

- 일반 진입 가구원과는 구별되며 목표 모집단의 변화를 나타냄.
- SOEP는 모형 방법으로 가구가중치가 그대로 유지됨.

○ 최근 이민자(recent immigration)

- 일반 진입 가구원과는 구별되며 목표 모집단의 변화를 나타내지만 SOEP에서는 구별하지 않음.

○ 사망(deaths)과 이민(emigration)

- SOEP는 모형 방법으로 사망자 가중치를 단순히 제거하며, 가구가중치는 그대로 유지함.

○ 합가(household mergers)

- 합가는 두 가지 경우로 볼 수 있음. 먼저 처음부터 서로 독립되어 있던 표본 가구원이 합치는 경우(unrelated merge)와 기존에는 동일한 가구에 있다가 분가한 뒤에 다시 합가하는 경우(move-back merge)임. 후자는 주로 학업을 마친 후 부모 가구로 들어오는 경우임.
- 모형 방법은 처음부터 서로 독립되어 있던 표본 가구원이 합치는 경우(unrelated merge) 일반적인 진입 가구원과 동일한 방법을 적용하는데 다만 추출확률이 다르다는 것을 명시함. 분가 후 다시 합가하는 경우는 가구가중치를 그대로 유지하는데 가구 추출확률이 변하지 않기 때문임.

○ 가구 분리(household splits)

- 자녀가 성인이 되어 분가하거나, 이혼하여 각각 가구를 분리하여 가지는 경우 등을 의미함.
- SOEP는 모형 방법으로 일반적인 진입 가구원 방법을 적용함.

제3절 조사 방식 및 관리²³⁾

- 독일의 SOEP는 대면면접을 통한 자기기입방식, 즉 CAPI와 PAPI를 혼합하여 사용하고 있음.
 - 1998년부터 CAPI시스템을 단계적으로 도입하기 시작하였으며 CAPI 조사의 효과를 파악하기 위해 연령이나 하위 표본 등으로 조사 대상을 구별하여 진행함.
 - CAPI 프로그램은 In2view임.
 - 한편 성공적인 조사를 위해 표본, 설문지의 종류에 따라 PAPI 방식을 혼합하여 사용함.
 - SOEP 조사에서는 조사 도구에 따라서 여러 가지 효과가 나타나는데 이것을 분석하기 위해 응답자 및 면접원의 반응에 대한 정보를 분석함.
 - 도구에 따른 조사 성공률에 대한 연구 결과 CAPI 사용으로 나타날 수 있는 응답자의 거부감, 즉 모드효과(Mode effect)는 거의 없는 것으로 나타남(Schupp, 2005; 이상호, 2006).
- 1996년 이후 표본 집단에 따른 설문지 차이가 없어지고 데이터 수집 또한 통합적으로 이뤄지고 있음.
 - 1996년 이전까지 SOEP는 다양한 표본으로 구성되어 각각의 표본에 따라 상이한 설문지가 설계되어 있었음.
 - 초기 만 16세 이상의 독일 거주자를 대상으로 매해 가구 조사를 실시함. 기존의 표본에 포함되어 있는 개인과 그 자녀, 중간에 표

23) Anne B, et al(2018)에 수록된 조사현황 자료를 중심으로 기술함.

본에서 이탈한 대상자들도 추적함.

- 1999년 청소년과 아동에 대한 설문지가 도입되어 전 가구 구성원을 대상으로 설문을 실시하고 있음.

□ 표본 집단에 따른 설문지 차이는 없지만 가구 구성원에 대한 측정오차를 줄이기 위해 가구 설문지와 가구 구성원 개인 설문지를 구분하고 있음.

- 가구 설문지의 응답자는 가구주이며 원표본 가구와 신규 표본 가구를 구분하여 조사함. 설문지의 첫 부분은 가구의 이름과 가족 구성 등 가구의 일반적인 정보를 기록함.

- 가구 구성원 개인 설문지는 기존 가구원용과 신규 가구원용이 구분되어 있음.

□ SOEP 조사는 모든 설문 정보를 가구원 본인에게 직접 얻고자 함.

- 타 국가의 주요한 가구조사 패널에(PSID, SIPP 등) 가구주에게 다른 가구원에 대한 정보를 얻는 대리응답이 사용된다는 점과 비교할 때 차별적임.

- 앞에서 살펴본 바와 같이 조사 이전의 응답자 접촉 방식에 대해서는 최대한 유연한 기준을 적용하고 있음.

□ SOEP 조사에서 조사원은 1인이 1구역을 맡아 관리함(이상호, 2006).

- SOEP 조사는 가능한 한 조사 지역을 세분하여 구역(cluster)으로 설정하고 조사원 1인에게 1개의 구역을 맡기는 방식으로 진행됨.

- 동일한 조사원을 매년 지속적으로 유지함으로써 조사의 안정성을 높이고자 함. 그러나 담당 조사원이 이탈하게 될 경우 해당 구역 조사의 성공에 상당한 영향을 미칠 수 있다는 점이 유의됨.
 - 1987년 4차 웨이브 이후 패널 대상자 관리를 위해 현금을 지급하는 대신 5~10마르크 상당의 선물(로터리 티켓)을 지급하고 있음.
- SOEP는 매해 조사 과정에 대한 정보 또한 누적하여 메타 데이터 형식으로 저장하고 있음.
- 모든 조사 연도에 한 번이라도 면접 조사를 받았거나 조사 대상이 되었던 응답자들의 조사 과정에 관한 기록 등이 담겨 있음.
 - 조사 과정에 대한 메타 데이터로 투입 인원, 조사 방식의 효율성, 조사 방식에 따른 조사시간의 차이 등 조사 전반을 관리하고 있음.
- SOEP 조사에 투입된 인원을 개관하면 다음과 같음.
- SOEP 조사는 1984년부터 가장 기본적인 원칙으로 현장에서의 면대면 방식을 지키도록 하고 있음. 이는 2000년 이래 PAPI로 실현함. 이후 2011년 J 유형의 집단 조사부터는 컴퓨터를 활용하는 CAPI 방식을 시작함(J, H, L1 조사는 CAPI 방식임).
 - 현장조사를 지휘하는 TNS 인프라테스트가 투입한 조사원의 규모를 조사 집단 유형별로 구분하여 보고하고 있음. 이는 다음의 <표 4-4>와 같음.

〈표 4-4〉 인터뷰 조사원 투입 규모(2016년)

(단위: 명, 가구)

구 분	A~H집단	J/K집단	L1집단	A~L1집단
인원수	432	279	170	506
1인당 평균 방문 가구 수	13	12	7	20

자료: Anne B, et al(2018), p.26.

○ A부터 L1까지 전체 집단의 조사를 위해 총 506명의 인원이 투입되었음. A부터 H집단까지의 조사를 위해서는 432명, J/K집단을 위한 조사에는 279명, L1집단의 조사에는 170명이 투입되었음. 전체 투입 인원이 각 집단에 투입된 인원보다 적은 이유는 집단별로 조사 시기가 달라 중복으로 조사를 진행한 인원이 있기 때문임. 2016년 조사원 1인이 담당할 평균 가구 수는 20가구로 나타남.

□ SOEP 조사 과정의 주요 결과는 다음의 〈표 4-5〉와 같음. 이는 A~H 표본 집단과 최근에 생성된 J, K, L1 표본 집단의 경우로 구분해 개관함. J, K, L1 표본 집단의 설문은 중앙 일괄처리 과정이 없음.

〈표 4-5〉 2016년 SOEP 조사에서 A~H 표본 집단의 조사 과정 결과

구분	A~H		J		K		L1	
	가구	%	가구	%	가구	%	가구	%
(1) 표본 집단 수	7,547	100.0	2,138	100.0	1,218	100.0	1,291	100.0
- 계속 참여 가구 수	6,988	92.6	1,983	92.8	1,108	91.0	1,184	91.7
- 이탈 수	354	4.7	106	5.0	76	6.2	84	6.5
- 신규 수	205	2.7	49	2.3	34	2.8	23	1.8
(2) 조사 대상 가구 수								
조사하지 않음(사전 거절/이주/사망)	112	1.5	8	0.4	5	0.4	1	0.1
면대면 조사(PAPI, CAPI)	5,448	72.2	2,130	99.6	1,213	99.6	1,290	99.9
- 계속 참여 가구 수	5,236	69.4	1,975	92.4	1,103	90.6	1,183	91.6
- 이탈 수	86	1.1	106	5.0	76	6.2	84	6.5
- 신규 수	126	1.7	49	2.3	34	2.8	23	1.8
중앙 일괄처리(우편 조사)	1,987	26.3	-	-	-	-	-	-
- 계속 참여 가구 수	1,573	20.8	-	-	-	-	-	-
- 이탈 수	268	3.6	-	-	-	-	-	-
- 면대면 조사 이후 우편 조사	67	0.9	-	-	-	-	-	-
- 신규 수	79	1.0	-	-	-	-	-	-
(3) 응답가구 수								
면대면 조사(PAPI, CAPI)	5,004	91.9	1,883	88.1	1,046	85.9	1,122	86.9
- 계속 참여 가구 수	4,895	93.5	1,819	92.1	1,004	91.0	1,074	90.8
- 이탈 수	38	44.2	39	36.8	28	36.8	39	46.4
- 신규 수	71	56.3	25	51.0	14	41.2	9	39.1
중앙 일괄처리(우편 조사)	1,490	75.0	-	-	-	-	-	-
- 계속 참여 가구 수	1,392	88.5	-	-	-	-	-	-
- 이탈 수	55	20.5	-	-	-	-	-	-
- 면대면 조사 이후 우편 조사	17	25.4	-	-	-	-	-	-
- 신규 수	26	32.9	-	-	-	-	-	-

자료: Anne B, et al(2018), pp.28-29.

- 2016년 SOEP 표본 집단의 규모는 총 1만 2,194가구로 이 중 A~H 표본은 7,547가구, J, K, L1 표본은 4,647가구임. 이들 전체 가구 중 계속 참여 가구는 92.4%, 새롭게 추가된 가구는 2.6%, 이탈한 가구는 5.1%로 집계됨.
- 2016년 SOEP 표본 집단의 가구 중 조사 대상 선정 작업을 실시한 결과, A~H 표본 집단의 72.2%인 5,448명이 면대면 조사 대상자로, 26.3%인 1,987명이 중앙에서 관리하는 우편조사 대상자로 나타남. 나머지 1.5%는 사전 거절, 이주, 사망의 이유로 조사 대상자에서 제외함.
- 2016년 SOEP 조사 대상자 중 최종적으로 면대면 조사를 실시한 가구의 비중은 91.9%(A~H 표본 집단)임. 지난 조사 대비 계속 참여 가구의 비중은 93.5%로 나타남. 계속 참여 가구의 비중이 가장 높았던 수치 94.4%에는 다소 못 미치는 수치임.
- 2016년 SOEP 면대면 조사의 응답가구 비율은 91.9%(A~H 표본 집단)로 나타남. 지난 조사 대비 계속 참여 가구의 비중은 93.5%로 나타남. 이탈한 가구는 38가구로 지난 조사 대비 44.2%로 나타남.
- 2016년 SOEP J, K 표본 집단의 최종 응답가구 중 계속 참여 가구의 비중은 각각 92.1%, 91.0%로 집계됨. A~H 표본 집단의 경우가 93.5%인 것과 비교해 낮은 수치임. 코호트 조사로 일곱 번째 조사인 L1 표본 집단의 응답가구 중 계속 참여 가구의 비중은 90.8%로 나타남.

□ SOEP 조사는 PAPI, CAPI, 자기응답 방식으로 현장에서 진행되거

나 우편으로 진행되어 중앙에서 일괄 처리될 수 있음. 이를 개관하면 다음과 같음.

- A부터 H까지의 표본 집단의 조사 방식은 면대면 조사를 원칙으로 하여 혼합 방식(mixed Ansatz)을 선택하고 있음. 즉 PAPI, CAPI, 자기응답 방식, 우편조사 모두 가능함. 이전까지는 PAPI, CAPI만 가능하였고 자기응답 방식이나 우편조사는 예외적인 경우로 여겨졌으나 다년간의 경험으로 PAPI, CAPI와 통계적으로 유의미한 차이를 보이지 않는 것으로 나타나 혼합 방식을 사용하게 되었음. 조사원이 가구원 여러 명의 조사를 함께 진행할 경우 스스로 응답지에 체크하는 자기응답 방식이 가능해졌으며, 당시 집에 있지 않은 가족 구성원의 경우 우편으로 설문지를 중앙에 보내는 방식이 허용되게 됨.
- 비교적 최근 도입된 J/K, L1 표본 집단의 경우는 A~H 표본 집단과 다르게 CAPI 방식만을 이용함. 표집을 도입할 당시 면대면 인터뷰의 장점을 최대한 살리면서 보다 효과적으로 중앙에서 일괄 작업하는 것을 위하여 CAPI 방식으로만 진행하도록 결정됨.
- 다음의 <표 4-6>은 위에 언급한 네 가지 조사 방식의 비율을 보여 주고 있음.

〈표 4-6〉 조사 집단별 인터뷰 방식의 비중(2016년)

(단위: %)

표본	현장 인터뷰 방식				중앙 일괄처리
	CAPI	PAPI	자기 응답	전체	우편
A~D	25.4	12.2	35.8	73.4	26.6
E	-	-	-	-	100.0
F	36.1	14.4	31.4	81.9	18.1
G	34.0	8.3	40.5	82.8	17.2
H	64.4	3.2	23.1	90.7	9.3
A-H	33.2	11.5	33.3	78.0	22.0
J/K	100.0	-	-	100.0	-
L1	100.0	-	-	100.0	-
전체	58.7	7.1	20.6	86.4	13.6

주: 가구 구성원 개인 조사.

자료: Anne B, et al(2018), p.31.

- 2016년 조사에서 A~H집단의 44.7%에 대해 직접 대면하여 설명하면서 조사가 진행되는 면대면 방식으로 조사가 실시되었음. 그리고 전체의 78%는 종이 설문지를 보며 직접 응답 및 기입 하였으며 22%에 대해서는 우편으로 조사함.
- 전반적으로 오래된 표집 대상의 조사일 경우 우편 조사의 비중이 높아지는 것으로 나타남. 이는 표집 집단이 오래되었을 경우 이미 여러 번 조사를 받은 대상자가 많아 중앙에서 관리하는 것이 더 효율적이기 때문인 것으로 해석됨.
- A~D 표집집단의 조사에서의 자기 응답 비율은 35.8%로 그보다 최근의 F, H 집단에 대한 조사에서보다 높게 나타남. 아울러 소득 상위계층에 대한 조사인 G 표집 집단의 자기 응답 비중은 40.5%로 가장 높게 나타남. 이는 소득과 자산에 대한 질문의 비중이 높기 때문으로 해석됨.

- SOEP 패널조사에서 가구 조사와 가구원 개인 조사 시간은 PAPI 방식과 CAPI 방식일 경우 차이가 나타남.
- 가이드라인에서는 설문조사를 위한 인터뷰 시간을 가구 조사의 경우 15분, 가구원 개인 조사의 경우 30분으로 제시하고 있음. 아울러 평균 가구원 수가 2명임을 감안하여 거주지에 방문할 때 인터뷰 시간은 총 75분으로 잡고 있음(여기에는 주소지 기입과 추가 설문지 질문 시간은 포함되어 있지 않음).
- 현장에서의 인터뷰 시간을 관리하기 위해서 인터뷰가 끝나고 조사에 걸린 시간을 기입하도록 함. 이를 기반으로 하여 PAPI 조사 방식과 CAPI 조사 방식의 인터뷰 시간을 계산함. A부터 H 집단까지의 인터뷰는 PAPI와 CAPI를 혼용하고 있어 비교가 가능함. 그러나 J, K, L1 집단의 인터뷰는 CAPI 방식으로만 하게 되어 있어 비교하지 못함. 다음의 <표 4-7>은 SOEP 2016 패널조사의 2015년과 2016년 평균 인터뷰 시간을 보여 줌.

〈표 4-7〉 PAPI와 CAPI에 평균적으로 소요되는 조사시간

구분	가구조사 시간				가구 구성원 개인 조사 시간				방문 가구당 소요 시간			
지침	15분				30분				75분			
연도	2015		2016		2015		2016		2015		2016	
조사 방식	PAPI	CAPI	PAPI	CAPI	PAPI	CAPI	PAPI	CAPI	PAPI	CAPI	PAPI	CAPI
A	18	16	18	15	36	33	41	45	90	82	100	105
B	17	16	18	14	38	32	39	42	93	80	96	98
C	26	17	27	16	43	34	48	46	112	85	123	108
D	19	16	18	16	40	34	48	47	99	84	114	110
E	-	-	-	-	-	-	-	-	-	-	-	-
F	22	17	23	16	39	35	44	48	100	87	111	112
G	20	16	20	15	37	33	42	44	94	82	104	103
H	22	17	25	17	39	35	50	46	100	87	125	109
J	-	17	-	16	-	36	-	48	-	89	-	112
K	-	16	-	16	-	35	-	47	-	86	-	110
L1	-	19	-	18	-	36	-	47	-	91	-	112
전체 (A~H)	21	16	22	16	39	35	44	47	99	86	110	110
전체 (A~L1)	-	16	-	16	39	35	-	47	-	86	-	110
지침과의 차이	+6	+1	+7	+1	+9	+5	+14	+17	+24	+11	+35	+35

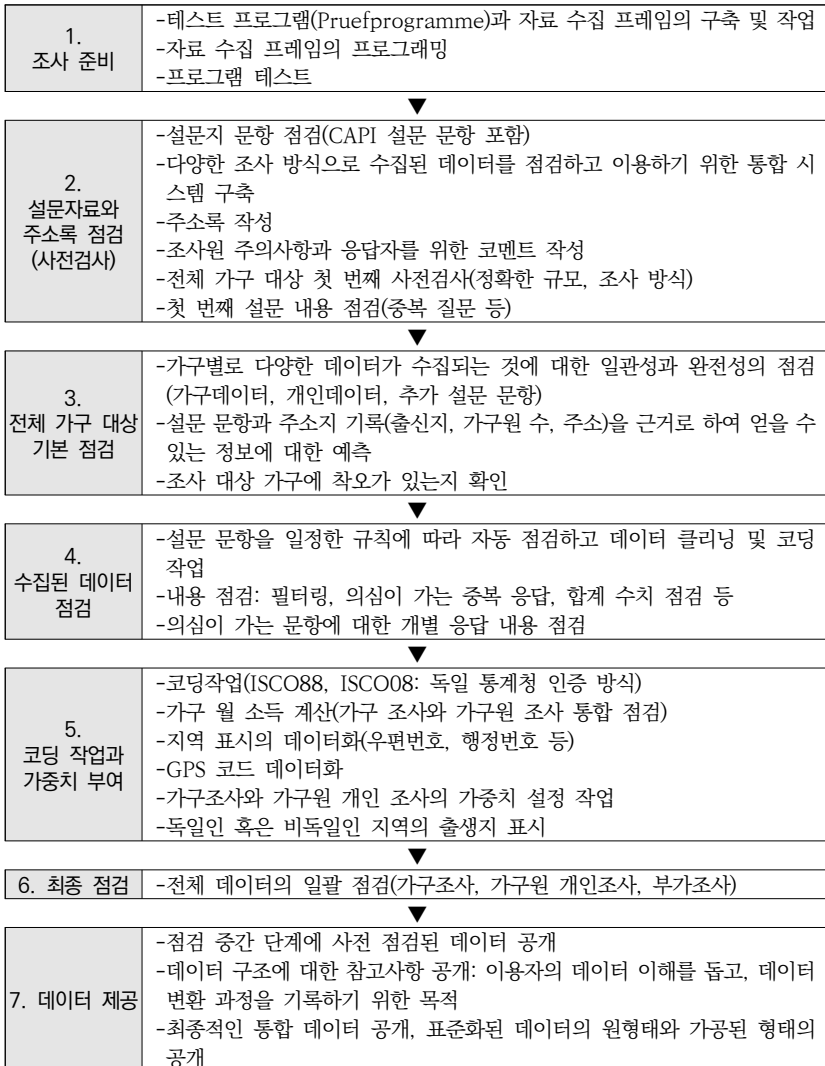
자료: Anne B, et al(2018), p.16.

○ 2015년 방문 가구당 평균 인터뷰 소요 시간은 PAPI와 CAPI 방식에서 각각 99분, 86분으로 나타남. PAPI 방식의 경우 지침(75분)보다 24분 더 걸렸고 PAPI 방식의 경우 지침보다 11분 더 걸린 것임(주소 기입 시간과 부가설문지 조사 시간은 제외함).

○ 2016년 가구 전체 조사의 평균 인터뷰 시간은 PAPI와 CAPI 방식을 모두 통합해 110분으로 나타남. PAPI 방식의 경우 지침(75분)보다 35분 더 걸렸고 PAPI 방식의 경우 또한 동일하게 35분 더 걸린 것으로 나타남.

□ SOEP 패널은 데이터 품질 관리를 위해 7개의 데이터 관리 단계를 수립하고 있음. 이를 개관하면 다음의 [그림 4-2]와 같음.

[그림 4-2] SOEP 패널의 데이터 수집 및 관리의 7단계



자료: Anne B, et al(2016), p.47.

제4절 조사표의 주요 내용

□ SOEP의 설문지는 가구용, 가구 구성원 개인용으로 구분되어 있음.

○ SOEP는 독일 인구 및 소득 통계의 수립을 목적으로 함. 고용, 교육, 자산, 소득, 개인이 가지고 있는 정부 정책에 대한 가치관 등을 대해 조사하고 있음.

○ SOEP 조사 초기의 설문지별 주요 질문 영역을 살펴보면 다음의 <표 4-8>과 같음.

<표 4-8> 가구 설문지 주요 문항

1. 가구 구성	- 가구 구성원의 변화
2. 주거 구성	- 주거지 형태 및 규모 - 주택 소유 형태
3. 이웃	- 이웃과의 교류 정도 - 주변 환경 평가
4. 주거 비용	- 세입자용: 월세, 주거비용 보조 여부 - 자가소유자용: 자가취득 방법, 주택 관련 지출
5. 가구소득, 자산, 이전소득	- 정부 보조금, 월 가구 순소득, 자산

자료: 강석훈(1997), p.7. 재인용.

<표 4-9> 가구 구성원 개인 설문지 주요 문항

1. 노동 참여자 - 현재 고용 상태 - 과거 1년간의 고용	6. 지난해 동안의 노동소득 및 이전소득
2. 직업이 있는 사람 -현재의 직업에 관한 사항	7. 은퇴 및 연금
3. 실직·실업·구직자 -가장 최근의 직장 -미래의 취업계획	8. 세금
4. 향후 5년간 일하길 원하는 사람 -원하는 주간 근무시간 수	9. 정부 및 기타 보조금
5. 50세 이상이며 직업이 있는 사람 -직업경력	10. 교육 및 직업훈련
	11. 시간활용
	12. 건강
	13. 삶의 만족도
	14. 정치적 선호
	15. 사회보장
	16. 사회적 배경
	17. 인구학적 특성

자료: 강석훈(1997), p.8. 재인용.

〈표 4-10〉 비독일인 응답자 주요 질문 문항

1. 교육	- 독일에서의 교육 - 외국에서의 교육
2. 지난해 동안의 이전소득	- 친척으로부터의 보조금 - 고국으로의 송금액
3. 인구학적 특성	- 출생국가, 배우자의 영주, - 독일 이주 연도, 희망 체류기간
4. 사회적 배경	- 부모의 영주, 부모의 연령 및 학력
5. 주관적 질문	- 국민ID번호 - 자신이 판단하는 언어 구사 정도 - 독일인과의 접촉 정도 - 현재 친구의 범위에 있는 사람들의 국적

자료: 강석훈(1997), p.8, 재인용.

□ 2016년 SOEP 조사의 가구용 설문지 질문 문항을 개관하면 다음과 같음.

- 주거지와 주거비에 관한 질문
- 정부보조금, 임대비, 이자소득 등 소득에 대한 질문
- 대출납입금 혹은 적금 등의 자산 형성을 고려한 수입에 관한 질문
- 돌봄을 필요로 하는 동거 가구원, 돌봄을 필요로 하는 정도, 주 돌봄자에 관한 질문
- 유치원, 학교 등에 재학 중인 아동에 관한 질문

□ 첫 번째 영역은 주거지와 주거비용(Wohnkosten)에 관한 질문임.

- 문항 1)은 주거지에 대한 내용으로, 이전 설문 당시와 동일한 곳에서 주거하고 있는지 아닌지, 주거지가 바뀌었다면 이전에 살던 곳과의 차이점은 무엇인지, 주거인의 변화는 있었는지 등에 관한 질문. 문항 6)은 주거지 건물의 건축 연도에 대한 질문임.

- 문항 7)에서는 어르신이 있는 경우 어디에 거주하고 있는지, 그리고 요양원에 거주하는지를 질문함.
 - 문항 18), 19)에서는 지난 한 해 동안 거주지 혹은 근처의 땅, 시설 등을 사용하기 위해 지출한 액수, 문항 20), 21)에서는 지난 한 해 동안 살고 있는 집의 개조나 유지를 위해 비용을 지출했는지, 만일 지출했다면 액수는 어느 수준인지를 질문함.
 - 문항 22)는 2016년에 새롭게 추가되었는데 지난 한 해 동안 현재 살고 있는 거주지를 유지하기 위해 지출한 기본적인 세금(Grundsteuer)에 관한 질문으로, 문항 23), 25)는 기본적인 세금 외에 난방, 전기 등 거주지 유지를 위한 부가적인 지출(Nebenkosten)에 관한 질문으로 구성됨.
 - 문항 31)에서는 정부 임대아파트(Sozialwohnung)에 살고 있는지를 질문함. 문항 33)에서는 거주지가 사유건물인 경우 주택조합 건물인지 지방자치단체 건물(kommunalen Wohnung)인지를 물어보고 있음. 문항 41)에서는 거주지 임대료에 대한 경제적인 부담을 질문함.
 - 문항 43)에서는 현재 거주지의 주소와 우편번호를 기입하도록 함.
- 두 번째 영역은 가구의 소득과 지출에 관한 질문임.
- 문항 46)에서는 지난 한 해 동안의 거주지와 땅 임대료 수입의 가능한 한 정확한 총액수를 기입하도록 함. 문항 47), 49)는 2016년에 새롭게 추가된 문항으로, 임대한 건물 혹은 땅의 대출 상환금, 이자, 혹은 개보수 비용을 질문함. 문항 48), 50), 52)는 이전과 동일한 질문으로 재작년의 상황을 질문함.

- 문항 56), 61)에서는 지난 한 해 동안의 유산, 사례, 로또 등의 수입을 질문. 문항 62), 63)은 2016년에 새롭게 추가되었는데, 주식으로 인한 수입과 그로 인한 세금 등을 질문함. 이어서 문항 64)에서는 이익 배분(Gewinnausschüttung)이 앞 질문에 포함되어 있는지를 질문함.
- 문항 65)는 정부에서 지원받은 금액에 대한 질문으로 지난 한 해 동안의 총액수를 질문함. 이어 문항 66), 67)은 지난 한 해 동안 지원이 지속되었는지 아니면 중간에 지원이 없어졌는지에 대한 내용임. 문항 70)에서는 정부지원금에 대해 지난 한 달 동안의 총액수를 질문하고 있음. 문항 71)에서는 정부지원금의 유형을 질문함.
- 문항 74)에서는 가정 내 물품에 대해 질문함. 컬러텔레비전, 유선 전화, 휴대전화 등이 거주지에 있는지를 질문함. 아울러 가족의 활동과 거주 환경에 관한 종합 정보를 다섯 가지 영역으로 구분하여 질문함.
 - 적어도 한 달에 한 번 영화 관람, 콘서트 관람, 스포츠 활동 등의 여가활동을 가족원과 함께하는지에 대한 정보
 - 낡은 의류 몇 개를 새로 구입해 교체했는지에 대한 정보
 - 거주 공간을 동절기에 따뜻하게 유지했는지에 대한 정보
 - 가구원 모두가 현재 자기에게 맞는 사이즈의 신발을 두 켤레 이상 갖고 있는지, 계절별 신발을 구비하고 있는지 등에 대한 정보
- 문항 77)에서는 거주지 내 애완동물(가축 제외)에 관해 질문함.

□ 세 번째 영역은 돌봄이 필요한(pflege-oder hilfebeduerfigen) 동거인에 관한 질문임.

○ 2016년에 새롭게 추가된 영역으로 문항 79)는 앞의 문항 67) 정부 지원금을 계속해서 받고 있는지를 묻는 것과 관련하여 돌봄이 필요한 동거인이 있는지를 질문하고 있음. 문항 80)에서 문항85)까지는 이와 관련하여 보다 자세한 정보, 예컨대 돌봄이 필요한 동거인이 몇 명인지, 그들 각자는 어떻게 일상을 유지하고 있는지, 주 돌봄자가 누구인지 등을 질문하고 있음.

□ 네 번째 영역은 동거 중인 아동(Kinder)에 관한 질문임.

○ 문항 89)에서는 동거하고 있는 아동이 학교에 다니고 있는지 여부를 질문함. 동시에 문항 90)에서는 학교에 다니는 경우 어느 단계의 학교에 다니고 있는지를 구체적으로 질문함.

○ 문항 92)에서는 학교에서 아동은 수업 외에 주로 어떤 활동을 하는지 질문함. 문항 96)은 아동의 돌봄(Betreuung)으로 어떤 활동을 하고 있는지 질문함. 유아원(kita) 등. 그리고 그 비용도 묻고 있음. 문항 97)은 2016년에 처음 도입된 문항으로 아동이 얻고 있는 소득에 대해 질문함.

□ SOEP의 가구원 개인용 설문지 질문을 개관하고자 함.

○ 2016년 가구원 개인에 대한 설문지가 대폭 수정되었음. 질문 영역, 즉 설문의 주제는 그대로 유지하되 현재 변화된 독일의 상황을 반영하여 개별 질문을 수정하였음.

□ 첫 번째 영역은 전반적인 생활에 관한 질문임.

- 1985년부터 문항 1)에서는 응답자의 전반적인 생활만족도를 묻고 있음. 11척도, 즉 0(전혀 만족스럽지 않음)부터 10(완전히 만족함)까지로 측정함. 이 외에 건강, 수입 등에 대한 전반적인 만족도를 묻는 개별 문항은 축약되었음. 대신 2016년에 가족 및 주변 인과의 관계에서의 만족도, 민주주의에 대한 만족도 등을 추가함.
 - 2016년 도입된 문항 3)에서는 다각도의 어려움의 빈도를 질문함. 즉 지난 2주 동안의 불안, 우울, 희망 상실의 빈도를 측정함. 문항 6)에서는 2012년 이후부터 지속된 ‘인생에서 중요하게 생각하는 것들’에 대한 질문을 믿음과 종교에 대한 내용으로 대체함. 아울러 문항 7), 8)은 ‘돈에 대한 생각’을 질문하던 것을 ‘통장 잔액을 봤을 때의 생각’으로 대체함. 문항 9)는 2014년부터 시작되었는데, 일주일 중 평균적인 근무일을 고려하여 근무일에 하는 일들이 무엇인지 질문하고 있음.
- 두 번째 영역에서는 지난해의 경제활동 상황을 질문하고 있음.
- 지난 한 해 동안의 응답자의 경제활동에 대해 현재와 비교하여 질문함.
- 세 번째 영역은 현재 가구원 개인의 경제활동에 대한 질문임.
- 문항 51)은 2014년 도입된 질문으로 응답자가 종사하는 직업에서 요구하는 작업이 무엇인지에 대한 정보를 수집함. 문항 59), 60), 61)은 2011년 도입된 질문으로 응답자의 월 급여와 그가 현재 다니는 회사의 이사회 등에 미칠 수 있는 영향력을 어떻게 생각하는지에 대해 질문하고 있음.

○ 문항 66)은 2011년 도입된 질문으로 응답자가 느끼는 자신의 일에 대한 중압감과 압박을 4점 척도로 질문함. 문항 85)에서는 응답자가 회사차를 개인적으로 사용할 수 있는지 질문하면서 문항 86)에서는 그 사용가치를 금액으로 환산하면 얼마가 될지 질문하고 있음.

○ 문항 91), 92), 98)에서는 조세 6등급(Steuerklasse VI)과 관련하여 부업을 몇 개나 하고 있는지 질문하고 있음. 문항 99), 100)에서는 지난 한 달간의 수입과 지난 한 해 동안의 수입을 구별하여 질문하고 있음.

□ 세 번째 영역은 건강과 질병에 관한 질문으로 구성됨.

○ 2014년 도입된 질문 영역임. 문항 106), 107)은 일상생활에서의 건강 상태에 대한 질문으로 구성됨. 계단 오르거나 다른 일상에서 겪는 장애 상황에서 활동의 어려움의 정도를 질문함. 문항 108)에서는 지난 4주간의 마음의 상태, 예컨대 실패감 등에 대해 질문하고 있음.

○ 문항 112), 113)은 생활습관에 대한 질문으로 흡연 빈도, 예컨대 담배, 파이프, 시가 습관에 대해 질문함. 문항 114)에서는 E-담배 습관에 대해 질문함. 아울러 문항 115), 116)에서는 음주습관에 대해 질문하고 있으며 문항 117), 118)에서는 육류 섭취와 야채 섭취 습관에 대해 질문함.

○ 문항 119), 120)은 신체에 관한 질문으로 신장과 몸무게 등을 조사함. 문항 121)은 만성질환에 대한 질문으로 최소한 일 년 이상 통증을 느낀 상황에 대해 조사하고 있음.

- 문항 128), 129), 134)는 2016년에 처음 도입된 질문으로 민간 수발보험과 개인적으로 지출하고 있는 건강 관련 비용을 조사함. 아울러 문항 135), 136), 137)에서는 개인적으로 가입한 민간 질병보험에 대해 질문하고 있음. 또한 문항 138)은 보험회사 명칭, 문항 139)는 매월 지급하는 보험금 액수를 질문하고 있음. 그 밖에 문항 140)에서는 가족 중 보험에 가입한 수를 조사하고 있으며 문항 141), 142)에서는 가능한 공제 수준에 대해 질문함.

□ 네 번째 영역은 주관적인 인식에 대한 질문으로 구성됨.

- 문항 147)에서는 4점 척도로 개인의 전반적인 가치지향성을 질문하고 있음. 이주민에 대한 주관적인 인식을 묻는 문항으로 149), 150) 질문이 2016년에 도입됨. 이주민의 수용, 일자리에 대한 다양한 시각을 11점 척도(0~10)로 조사함.
- 문항 151), 152), 153)은 2011년에 도입된 질문으로 알고 지내는 사람, 친구, 가족에 대한 주관적인 생각을 조사하고 있음.

□ 다섯 번째 영역은 가족의 상황과 혈통에 대한 질문으로 구성됨.

- 문항 157)은 2016년에 도입된 질문으로 개인의 성적인 선호를 묻음.
- 독일 시민권을 갖지 않은 비독일인인지를 문항 163)에서 질문함. 이후 문항 164)~168)에서는 국적이 어디인지, 그리고 모국과 독일의 생활이 어떻게 연결되는지(소득이전 등)에 대해 질문하고 있음.
- 문항 171)~174)에서는 지난 10년간 각종 영역에서의 만족도를 질문하고 있음.

- 한편 SOEP 조사는 가구조사와 가구원 조사에 특별 주제를 포함하여 질문하고 있음. 1984년 이후의 특별주제를 개관하면 다음과 같음.
- 직업 경력, 혼인사 및 가족사, 사회적 배경, 직업 경로, 주거 환경, 사회보장, 은퇴, 자녀 양육, 자산 및 부채, 교육 및 직업훈련, 시간 활용 및 선호, 가족과 재정적 지원 등

제5절 패널 활용

- SOEP 데이터 관리는 설문조사를 총괄하는 TNS 인프라테스트에서 시행함.
- 1차적인 데이터 에디팅과 코딩, 클리닝 작업을 하고, 1차 데이터 클리닝을 마친 후 DIW와 프랑크푸르트, 만하임, 베를린 대학에서 다시 클리닝 작업을 하고 있음.
- SOEP 자료는 관계형 데이터베이스 시스템에 저장되고 있으며, 데이터베이스의 프로그램은 SIR을 사용함.
- 조사 자료는 조사 방식에 대한 정보, 가구 정보, 개인정보, 해마다 누적된 정보 등을 분류하여 보관함.
- SOEP 자료를 저장하는 SIR 관계형 데이터베이스 시스템에서는 분류되어 있는 모든 파일을 목적에 따라 조합해 추출해 낼 수 있음.
- 웹 기반의 'SOEPINFO'로 데이터를 추출함.
- SOEP 자료는 SIR-Export 파일, ASCII 파일, SAS/SPSS 파일로서 시디롬 형태로도 배포되고 있어 다양한 활용을 지원하고 있음.

- 독일 내의 각 대학과 연구기관, 정부 부처는 SOEP 자료를 독일 거주민의 행동과 독일 사회 변화에 대한 이해를 높이기 위한 다양한 분석과 연구에 활용하고 있음.
 - 또한 1991년부터는 국가 차원의 조사 자료를 해외에 배포하는 것을 금지했던 법이 개정되면서 미국, 영국, 캐나다, 네덜란드, 호주 등 많은 외국에서도 SOEP 자료 활용이 활발해지고 있음.
 - SOEPlit(<http://www.diw.de/english/sop/soeppub/soeplit/index.html>)에서 SOEP 자료를 활용한 국내외 보고서, 논문 등의 온라인 문헌 검색 가능.
- 독일 SOEP는 유럽 가구패널 컨소시엄(Consortium of Household Panels for European Socio-Economic Research: CHER)에 자료를 제공하고 있음.
- CHER은 유럽 각국의 가구패널자료를 비교 가능한 데이터베이스로 만들기 위해 시작됨. 유럽연합(EU) 14개국 외에도 스위스, 폴란드, 헝가리, 미국 등 총 18개국이 참여하고 있음.
 - 자료의 내용은 인구학적 정보, 교육, 건강, 경제활동 상태, 소득과 지출, 주거 등의 변수가 제공되며 SAS, SPSS, Stata, ASCII 등의 형태로 자료가 제공됨.
 - 이 외에도 MMM-데이터베이스(Multi-Master replication Manager for MySQL Database)를 운영하여 거시적 비교가 가능한 자료 즉 사회보장, 고용, 거시경제와 관련된 정보들이 제공됨.
- 독일 SOEP 자료는 CNEF로 활용되고 있음.

- CNEF는 미국의 PSID, 영국의 BHPS, 캐나다의 SLID, 독일의 SOEP 4개국의 가구패널 중 비교가 가능한 중요 변수들을 표준화된 형태로 제공하기 위해 시작되었음.
- 1982년 자료부터 활용이 가능하고 자료는 계속 업데이트됨. 자료는 SAS, SPSS, STATA, ASCII 등으로 제공되며, 가구와 개인 ID가 있어 원자료와도 매칭이 가능함.
- 주로 가구소득과 관련된 정보들이 제공되고 있고, 그 외에 주거, 교육, 건강 고용 등에 대한 정보도 제공됨.

제6절 복지패널에 주는 시사점

□ 패널조사 표본의 다양화 전략

- 독일 SOEP 패널은 표본 집단을 지속적으로 다양화하여 추가하고 있음. 현재 A에서 L까지의 표본집단이 구성됨. 그 배경에는 독일 통일과 유럽 통합이라는 역사적 흐름, 이주민의 증가로 인한 인구구성 변화, 패널조사 기간의 장기화에 따른 표본 수의 자연 감소가 있음.
- 사회 변화가 일어난 경우 그와 관련된 특정 집단, 이민자 집단 등을 표본으로 추가함. 패널 장기화에 따른 자연스러운 표본 수 감소의 경우에는 전체 표본 구성을 재설정하는 것이 보통. SOEP는 두 경우 모두 표본 집단을 추가하는 방식으로 표본 수를 유지하면서 사회 변화의 특성을 감안하고 있음.
- 우리나라 또한 최근 이주민(다문화가정, 탈북자 등)의 증가, 농어

촌 인구 감소 등의 사회적 변화를 겪고 있음. 이로 인한 인구 구조의 변화를 보다 심층적으로 고려한 표본 설계 계획이 필요함.

□ 독일 통일 대비 표본 설계 방식 계획

- SOEP 패널은 독일의 공식적인 통일 시점 6개월 전부터 표본 집단 재설계 작업에 들어감. 그리고 동독 지역에 거주하는 독일인을 모집단으로 하여 1990년 새로운 형태의 표본 집단 C를 추가함.
- 단시간에 패널조사 표본 집단이 추가될 수 있었던 것은 이를 위해 전부터 동독 지역의 인구구조에 관심을 가졌으며 통일 시대 동서독 통합 패널의 가능성을 다각도로 모색해 왔기 때문임.

□ 사회 변화를 반영하기 위한 지속적인 설문 문항 개선

- SOEP의 설문 문항은 지속적으로 개선되고 있음. 그 방향은 인터넷을 통한 일상 활동의 증가, 전자로 통용되는 금전 등의 생활 트렌드 변화를 문항에 반영하는 것임.
- 또한 사회적 욕구의 변화를 반영하여 실천적, 정책적 연구의 함의를 높이는 문항으로 개선하고 있음. 예를 들어 가족 구성원 중 일상생활에서 타인의 도움을 필요로 하는 사람이 있는지, 그 수준은 어느 정도인지, 어떤 식으로 도움을 제공하고 있는지 등 케어 혹은 수발(pflege)에 관한 문항이 늘어나고 있음. 이는 노인인구 증가 등으로 독일 사회의 케어 욕구가 늘어나고 있는 상황을 파악하려는 의도를 가짐.

□ 가구 구성원 개인 설문지의 다양화

- SOEP 조사 설문 문항은 성인용, 청소년용(Jugend-Fragebogen),

초기 청소년용(Fruehe Jugend-Fragebogen) 그리고 재학 중인 청소년용(Schuelerinnen und Schueler-Fragebogen)으로 구분되어 있음. 성인용 설문 문항은 1984년 이후 그리고 청소년용 설문지는 1999년, 초기 청소년용 설문지는 2002년, 재학 중인 청소년용 설문지는 2004년에 도입되었음.

- 이러한 설문 문항의 가구원 특성에 따른 다양화는 독일 외 유럽 국가 가구조사패널의 일반적인 경향임. 이는 패널조사의 기본이 되는 설문 문항의 다양화를 의미함. 특별한 관심이 요구되는 계층에 대한 부가조사 설문 문항은 별도임.

□ 조사 방식으로 PAPI 대비 CAPI의 비중 확대

- 독일 SOEP 조사에서 컴퓨터 지원 면대면 인터뷰 방식의 비중이 증가하고 있음. 이는 CAPI 방식이 현장 인터뷰 방식, 데이터 수집 및 관리 등 조사 프로세스 전반에 단순화를 가져올 것으로 기대됨.
- SOEP 패널조사에 CAPI 방식의 도입이 PAPI, 면대면 자기 기입 방식, 중앙의 우편조사 방식과 비교하여 조사 프로세스 전반의 단축을 가져오는지 그리고 자료 수집의 정확성이 확보되는지에 관한 조사 프로세스 연구가 진행됨. 그리고 그 결과를 바탕으로 CAPI 방식을 점진적으로 확대하고 있음.

□ 조사 과정의 기록을 의무화하여 조사 프로세스 전반의 관리 시스템 구축

- 조사원과 응답자 관리를 통해 조사 응답률과 수집 자료의 정확성을 높이고자 함. 이를 위해 조사 과정 전반, 즉 조사 실시 기간, 조

사 방식, 면접 시간, 응답자의 태도 등을 기록하도록 하고 있음.

- 그리고 이러한 자료를 분석하여 공개하고, 궁극적으로 조사원과 응답자의 태도, 조사 방식, 지역효과 등이 조사에 미치는 비표본 오차를 최소화하기 위해 노력함.

□ 데이터의 다양한 활용을 가능하게 하는 자료 배포 및 분석 공유

- 다목적 조사의 목적을 실현하기 위해 이용자가 다방면의 분석에 데이터를 활용할 수 있도록 조사 주관 기관에서 데이터 클리닝, 가중치 등의 작업을 수행하기 전과 후 두 가지 형태의 데이터를 모두 제공함. 아울러 이용자의 편의를 높이기 위해 자료를 SAS, SPSS, STATA, ASCII 등으로 제공하고 있음.
- 또한 여러 방면에서 데이터를 분석하여 서로 교류하는 것이 데이터 활용의 시너지를 가져온다는 취지에서 SOEP 분석 연구를 SOEP 홈페이지에 등록하도록 독려함. 그리고 이를 정리하여 연구 영역별로 분류 공개하고 자체적인 워킹페이퍼 시리즈 발간, 학술대회 개최 등을 통해 조사의 이론적, 실증적 성과를 공유함. SOEP 패널 이용자의 데이터 활용을 지원하기 위해 데이터 수집 과정 및 구조에 대한 설명회 또한 주기적으로 개최함.

□ 국제적 교류의 활성화

- SOEP는 국제 패널 네트워크에서 중요한 역할을 하고 있음. 국가 비교 연구의 데이터베이스를 갖추기 위해 타 국가의 패널조사와 교류하고 있음. 주요한 데이터 연계 상황은 다음과 같음.
 - 횡단연구 데이터베이스: LIS, Luxembourg Wealth Study (LWS)

- 종단연구 데이터베이스: Cross-National Equivalent File(CNEF), Consortium of Household Panels for European Socio-Economic Research(CHER), European Community Household Panel(ECHP)

○ 이를 위한 기초 작업으로 국제화된 코딩 분류를 따르고자 노력함.
예를 들어 교육은 ISCED, DASMIN으로 코딩할 수 있도록, 노동 시장은 ISCO88, NACE로 코딩할 수 있도록, 지역정보는 NUTS로 코딩할 수 있도록, 연소득은 Canberra Group(2001)에서 추천한 구조로 코딩할 수 있도록 유연하게(flexible), 즉 국내 응답자에게 익숙하면서 국제화 코딩 분류로도 분석이 가능하게 문항을 설계하고자 노력함.

제 5 장

일본 Japan/Keio Household Panel Survey(JHPS/KHPS)

제1절 조사 배경

제2절 표본 및 가중치

제3절 조사 방식 및 관리

제4절 조사표의 주요 내용

제5절 패널 활용

제6절 복지패널에 주는 시사점

5

일본 Japan/Keio Household Panel Survey(JHPS/KHPS) <<

제1절 조사 배경

□ 조사 목적 및 경위

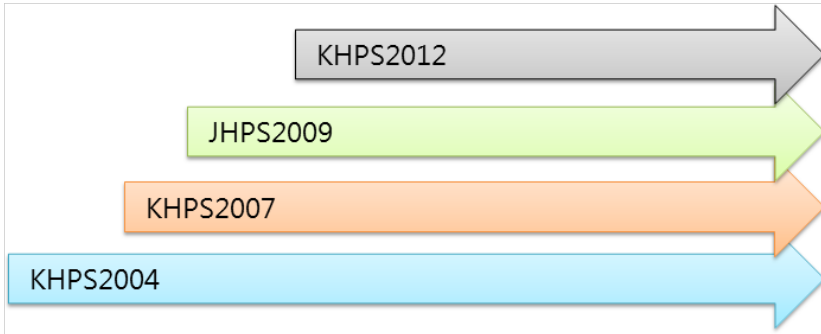
- 민주주의 국가에서 바람직한 제도·정책의 설계를 위해서는 신뢰할 만한 데이터와 그것을 이용한 정책평가 시스템이 필요함. 패널조사는 동일한 개인을 지속적으로 추적하여 경제 주체의 동학적인 행동 분석이나 관찰을 할 수 없는 이질성을 고려한 분석이 가능하다는 점에서 사회과학의 연구·정책평가에 매우 필요함.
- 그러나 일본에는 유럽(European Community Household Panel, ECHP), 미국 PSID와 같이 사회 전체의 인구 구성을 반영한 가계패널조사는 존재하지 않음.
- 이에 사회과학의 연구·정책평가의 분석, 신뢰할 만한 데이터에 따른 평가 시스템 구축을 목적으로 2004년 게이오가계패널조사(Keio Household Panel Survey: KHPS), 2009년 일본가계패널조사(Japan Household Panel Survey: JHPS)를 실시함.
 - 게이오가계패널조사(KHPS)
 - 사회과학에 대한 연구·정책평가 시 중요하게 된 경제 주체의 동학적 행동 분석 및 관찰을 목적으로 일본 사회 전체의 인구 구성을 반영한 가계패널조사.
 - 일본가계패널조사(JHPS)

- 국가에서 제도·정책 설계 시 중요하고 신뢰할 만한 데이터와 데이터를 이용한 정책평가 시스템 구축을 목적으로 연구성과의 발신, 공개를 통해 사회과학 분야의 제도·정책연구의 발전에 공헌하는 것을 목표로 조사 실시.

○ KHPS와 JHPS는 각각 개별 조사로 실시·관리되어 왔지만 2014년부터 두 개의 조사표를 통일하여 데이터 관리의 일원화를 도모하기로 함.

- 따라서 일본을 대표하는 가계패널조사라는 의미를 담아 ‘일본 가계패널조사(JHPS/KHPS)’로 명칭을 통일하기로 함.
- 게이오대학 패널데이터 연구센터에서 KHPS와 JHPS의 조사표나 이용신청·보고서를 통합하여 운용하고 있음.
- 이로 인해 두 데이터의 통합분석이 가능하게 되었고, 패널데이터에 있어 불가피한 샘플 탈락에 따른 표본 수 감소에 보조적으로 대처할 수 있게 됨.
- KHPS와 JHPS의 조사 항목은 비슷한 것이 많았기 때문에 조사표를 통일하는 과정에서 내용을 크게 변경하지 않아도 되는 장점이 있음.

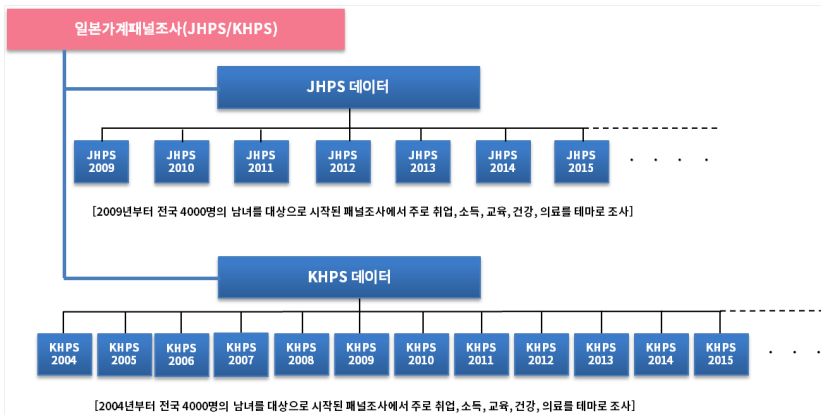
[그림 5-1] JHPS/KHPS의 4개 코호트



자료: 게이오대학 패널데이터 연구센터 홈페이지(데이터세트) <https://www.pdrc.keio.ac.jp/pandata/datasets/jhpskhps/#update>에서 2018. 5. 15. 인출

- 데이터의 구조와 샘플 등의 특성은 다음의 그림과 같음.

[그림 5-2] JHPS/KHPS의 구조



자료: 게이오대학 패널데이터 연구센터 홈페이지(데이터세트) <https://www.pdrc.keio.ac.jp/pandata/datasets/jhpskhps/#update>에서 2018. 5. 15. 인출

제2절 표본 및 가중치

1. 표본

○ KHPS

- 2004년부터 조사를 시작한 제1세대는 조사 시점에서 20세부터 69세까지의 남녀 약 4,000명을 대상으로 함.
- 표본 탈락에 대한 대응으로 2007년(제4회 조사) 조사 시점에서 20세부터 69세까지의 남녀 약 1,400명, 2012년(제9회 조사) 조사 시점에서 20세부터 69세의 남녀 약 1,000명을 추가함.

○ JHPS

- 2009년부터 조사를 시작한 제1세대는 조사 시점에서 20세 이상(연령 상한 설정 없음)의 남녀 약 4,000명을 대상으로 하고 있음.
- JHPS와 KHPS의 샘플 추출 모집단은 중복되지만, 조사 대상은 중복되지 않음.

○ JHPS와 KHPS의 차이점 및 특징

- JHPS는 조사 시작 이래 신규 코호트가 없지만, KHPS는 2007년 1,400명과 2012년 1,000명을 신규 코호트로 추가함.
- JHPS와 KHPS의 특징
 - 추출 단위는 개인 단위
 - 추출 방법은 층화2단 무작위 추출
 - 예정한 표본 수를 확보하기 위해 정규로 추출된 대상자(정규

대상자)와 예비 대상자[정규 대상자와 동일한 지역구, 동일 성별, 동일 연령대(열 살 단위)를 정규 대상자 1명당 10명]를 준비.

- 조사 대상자에게 배우자가 있는 경우 그 배우자에 대해서도 같은 조사를 실시함.

□ 표본 추출 방식

○ JHPS/KHPS의 조사 대상은 층화2단 무작위 추출 방법으로 선정.

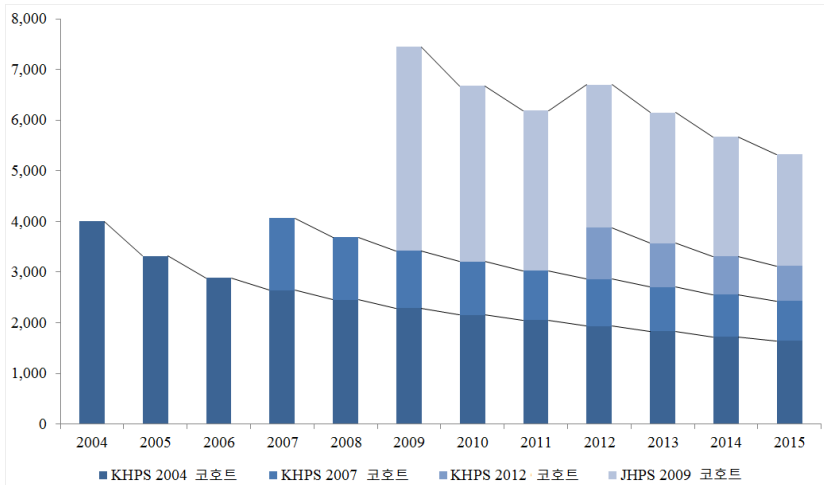
- 추출 제1단계에서는 전국을 지역·도시 계급으로 24층으로 층화하여 각 층에 주민기본대장(住民基本台帳)²⁴⁾의 인구 비율에 맞춰 표본 수를 배분.
 - 한 개의 조사지역당 표본 수를 10 정도(KHPS2007, KHPS2012는 5 정도)로 하여 각 층의 조사지역수를 결정하여 각층에서 소정의 조사구를 무작위로 추출.
 - 조사지역은 추출 단위로 ‘인구센서스’의 조사구를 사용하고 있음.
- 추출 제2 단계에서는 선정된 조사지역의 주민기본대장을 추출 대장으로 함.
 - 조사 대상 자격자를 대상으로 지정된 기번호(起番号)와 추출 간격에 근거하여 한 조사지역에 약 10명(KHPS2007, KHPS2012에서는 5명)을 추출함. 여기에서 추출된 대상자가 정규 대상자가 됨.

24) 주민기본대장은 성명, 생년월일, 성별, 주소 등이 기재되어 있는 주민표를 편성한 것임 (일본 총무성 홈페이지, <http://www.soumu.go.jp>).

- 단 정규 선정된 조사 대상자의 이사, 장기 부재, 주소 불명 등으로 만나지 못하거나 조사를 하지 못할 경우 미리 선정해 둔 예비 대상자로 대체하여 조사함.
- 상술한 바와 같이 예비 대상자는 정규 대상자 1명당 10명을 준비하는데 각각 정규 대상자와 같은 조사구에 거주하고 성별·연령도 같기 때문에 정규 대상자가 예비 대상자로 대체 되어도 지역·성별·연령 구분에서는 무작위 추출 결과가 유지된다고 할 수 있음.
- 조사 대상자에게 배우자가 있는 경우, 그 배우자에 대해서도 동일한 조사 항목을 사용하고 있음.

[그림 5-3] JHPS/KHPS의 표본 분포

(단위: 명)



자료: 게이오대학 패널데이터 연구센터(2016). Integrating the KHPS and JHPS to create the JHPS/KHPS(JHPS/KHPS User Guide), p.2.

2. 가중치

○ KHPS 및 JHPS의 조사 설계, 표본의 특성과 대표성, 회답지속경향은 각각의 조사에서 분석되어 왔음.

- 그러나 데이터 표본·바이어스에 대한 수정이나 표본 탈락에 따른 바이어스 수정에 대한 대책은 조사 시작 이래 강구되어 오지 않았음.

○ 최근 게이오대학 패널데이터 연구센터에서 바이어스에 대한 수정가중치, KHPS와 JHPS를 통합할 때 필요한 표본 바이어스를 수정한 가중치를 발표함(石井 加代子, 野崎 華世, 2014).

- 초기조사의 횡단가중치

- 초기조사 횡단가중치의 목적은 표본에서 발생하는 바이어스의 수정과 조사 대상자의 협력 거부, 즉 무응답에서 발생하는 바이어스의 수정임. 전자의 경우 일반적으로 선택확률의 역수인 이니셜(Initial) 가중치로 대처하지만, KHPS와 JHPS의 정규 대상자는 전술한 바와 같이 층화2단 무작위 추출로 지역 인구분포를 고려하여 무작위로 추출되었기 때문에 이니셜 가중치 작성은 생략함.
- 한편 예비 대상자의 예정 표본 수를 확보했더라도 얻어진 데이터에는 정규 대상자나 예비 대상자의 회답 거부로 발생하는 바이어스가 있기 때문에 이 부분에 대한 수정이 필요함.²⁵⁾ 이에 무응답 바이어스를 수정하기 위해 캘리브레이션

25) JHPS/KHPS에는 정규 대상자와 같은 거주지, 성별, 연령대의 예비 대상자가 있기 때문에 회답자를 예비 대상으로 대체하더라도 지역, 성별, 연령에서 무작위 추출의 결과를

(Calibration) 조정에 따른 방법을 사용.

- 캘리브레이션 조정은 가중치값이 모집단에 일치하도록 센서스 통계 등을 근거로 가중치를 조정하는 것으로, 레이킹(Raking) 조정(Iterative proportional fitting) 방법을 이용하여 학력, 취업 상태, 배우자 상황, 연령대별(5세 단위), 지역(8블록)의 변수를 축으로 하여 모집단으로 간주되는 벤치마크통계(주로 '인구센서스')에 맞춰 가중치 작성.

- 종단가중치

- 표본 탈락에 따른 바이어스를 수정하기 위해 초기조사의 횡단가중치를 바탕으로 각 웨이브의 가중치를 작성.
- 웨이브 1의 횡단가중치를 바탕으로 웨이브 1부터 5까지 회답지속인지 아닌지에 대해 웨이브 1 시점의 정보를 가지고 회답지속확률 모델을 추계.
- 추계된 회답지속확률의 역수를 이용하여 웨이브 2부터 5까지의 표본 탈락에 따른 바이어스를 수정하여 종단가중치를 작성.

$$W_{adjust, wave5} = W_{adjust, wave1} \frac{1}{P_{response, wave2}}$$

- $P_{response, wave2}$: 1차 웨이브부터 5차 웨이브까지의 회답지속확률
- 실제 분석 방법은 로짓모델(회답지속=1, 기타=0)을 이용해 회답지속확률을 추계하여 그 회답지속확률의 역수를 표본

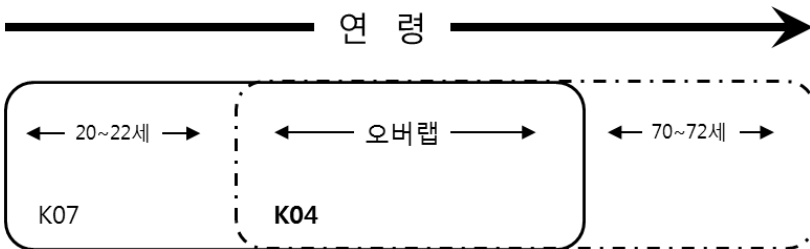
유지할 수 있음, 그러나 학력이나 소득 수준, 세대 유형 등 기타 조건은 통제할 수 없기 때문에 극단적인 예로 대상자 전체에서 고학력자의 조사 협조율이 0일 경우 얻어진 표본은 고학력자가 없는, 바이어스를 갖고 있는 데이터가 됨.

탈락에 따른 바이어스를 수정한 종단가중치로 함.

- JHPS와 KHPS의 통합 가중치

- 복수패널[KHPS2004표본(K04), KHPS2007 추가 표본(K07), KHPS2012 추가 표본(K12), JHPS2009(J09)]의 통합 방법을 검토. 단순한 방법으로서 각 샘플의 모집단 연령 분포 차이에 착목하여 통합 방법 검토.
- 예컨대 2007년에 추가된 K07을 K04와 통합한다고 하면 K04, K07 모두 표본 시점의 대상 연령이 20~69세로, 2007년 시점에 K04의 연령 폭은 23~72세, K07의 연령 폭은 20~69세임. 두 개의 표본이 겹치는 연령은 23~69세 부분으로 그 부분의 비중을 줄일 가중치를 부여함으로써 통합이 가능하게 됨.

[그림 5-4] KHPS2004 표본과 KHPS2007 추가 표본을 2007년 시점에서 통합



주: K07의 23~69세에 가중치 부여, K04의 23~69세에 가중치 부여.
 자료: 石井 加代子·野崎 華世(2014). p.12.

- 패널 통합용 가중치는 결합추정방법(Combining Estimates Method)을 사용. 이는 각각의 패널 표본 수와 샘플링 디자인 효과(design effect)에서 가중치를 작성하여 계산되는

가중치를 부여하고 각 패널을 모두 더해 모집단을 추계하는 방법을 사용하여 통합 가중치 산출.

제3절 조사 방식 및 관리

□ 조사 방식

○ 조사 빈도와 조사 방법

- 연 1회, 1월 31일 시점에서 자기기입식 조사 방법에 따른 조사를 함.
- KHPS는 전 대상자에게 자기기입식 방법에 따른 조사를 하지만 JHPS는 대상자에 따라 자기기입식 방법, 자기기입식과 면접(희망자는 인터넷상에서의 회답도 가능) 방법을 병행하여 조사함.
- JHPS는 조사 방법의 차이가 조사표 회수율에 미치는 영향을 밝히기 위해 실험적으로 다양한 방법을 사용함.
 - ① 두 개의 조사 방식 사용
 - ② 웹 회신의 옵션 설계
- 조사원의 동기 부여가 회수율에 미치는 영향을 보기 위해
 - ① 조사원 사전 설명 내용
 - ② 조사원 보수액의 차등 설계
- 2014년 이후부터 조사원이 조사 대상자의 자택에 방문하여 조사표를 주고 후일 재방문하여 조사 대상자가 기입한 조사표를 회수하는 자기기입식 유치조사법으로 통일.

○ 조사 실시 주체

- 조사 실시는 중앙조사회사에 위탁하고 있음.
- 중앙조사회사가 수집한 데이터를 분석하는 해석센터의 연구원이 주로 작업함.

○ 회수율

- 초기조사의 유의회수율(참고치)
 - 통상 추출된 대상자에 대해 조사 의뢰를 하고, 조사협력을 받지 못한 경우 무응답 처리가 되지만 KHPS와 JHPS는 전술한 바와 같이 소정의 표본 수를 확보하고 있기 때문에 정규 대상자에게 협력을 거부당한 경우 예비 대상자에게 협력을 의뢰함. 따라서 통상의 조사와 비교 가능한 회수율을 산출하기 어려움.
- 회답지속률
 - KHPS/JHPS 각각 제2회 조사 이후의 회답 상황과 조사회수율[(유효 회답 수-부활 표본)/전년도 완료 수×100]은 <표 5-2>, <표 5-3>과 같음.
 - KHPS와 JHPS 모두 조사가 거듭됨에 따라 회수율이 높아지고 있고, 대략 90% 전후의 회수율을 확보하고 있음.

〈표 5-1〉 초기조사의 유의회수율(참고치)

구 분	KHPS			JHPS
	2004년 1차 웨이브	2007년 추가샘플	2012년 추가샘플	2009년 1차 웨이브
조사 대상자 수	13,430	5,868	3,183	12,549
유효 회답 수	4,005	1,419	1,012	4,022
접촉 가능 대상 수	9,737*3	4,256	2,331	9,654
접촉가능률(%)	72.5*3	72.5	73.2	76.9
유의회수율(%)	41.1	33.3	43.4	41.7

주: 1) 유효 회답 수에는 예비 대상으로 조사한 사람도 포함.
 2) 접촉가능률(%)=(접촉 가능 수/조사 대상자 수)×100
 3) KHPS2004는 조사원 확인표를 도입하지 않았기 때문에 접촉가능률은 KHPS2007 추가샘플과 동일하다고 가정하여 접촉 가능 대상 수를 산출.
 4) 유의회수율=(유효 회답 수/접촉 가능 대상자 수)×100
 자료: 石井 加代子·野崎 華世(2014), p.5.

〈표 5-2〉 KHPS 회답지속률

차수(연도)	조사 대상자 수(명)	그중 전년도 완료 수(명)	유효 회답 수(명)	그중 복원샘플(명)	무응답(명)	회답 지속률(%)
2차 웨이브 (2005)	4,005	4,005	3,314	-	691	82.7
3차 웨이브	3,342	3,342	2,887	0	455	86.4
4차 웨이브	2,894	2,887	2,643	3	251	91.4
5차 웨이브	4,067	4,062	3,691	0	371	90.9
6차 웨이브	3,706	3,691	3,422	4	273	92.6
7차 웨이브	3,448	3,442	3,207	7	222	93.5
8차 웨이브	3,232	3,207	3,030	10	187	94.2
9차 웨이브	3,041	3,030	2,865	10	175	94.2
10차 웨이브 (2013)	3,088	3,877	3,568	11	320	91.7

주: (유효 회답 수-복원샘플)/전년도 완료 수×100
 자료: 石井 加代子·野崎 華世(2014), p.6.

〈표 5-3〉 JHPS 회답지속률

차수(연도)	조사 대상자 수(명)	그중 전년도 완료 수(명)	유효 회답 수(명)	그중 복원샘플(명)	무응답(명)	회답 지속률(%)
2차 웨이브 (2010)	4,022	-	3,470	-	522	86.3
3차 웨이브	3,490	3,407	3,160	6	316	90.9
4차 웨이브	3,170	3,160	2,821	4	342	89.1
5차 웨이브 (2013)	2,839	2,821	2,581	3	248	91.2

주: (유효 회답 수-복원샘플)/전년도 완료 수×100

자료: 石井 加代子·野崎 華世(2014), p.6.

○ 대상자 관리

- 인센티브

- 대상자에게 인센티브로 사례 3,000엔을 지급하고 있음. 사례를 지급하는 타이밍은 조사원이 조사표 회수 후에 집적 전달함.

- 대상자와의 연락

- 조사 실시 전, 대략 1월에 사전협력 의뢰장을 보내고 있음. 2월부터 3월에 걸쳐 조사를 의뢰하고 조사원은 조사표의 배포와 회수를 통해 대상자 1명당 적어도 2회 방문하고 있음. 7월에는 복중문안(暑中見舞)엽서, 12월에는 뉴스레터를 보내고 있음.

- 연락처 확보

- 사전협력 의뢰장, 뉴스레터 송부 시 주소 이전 연락 엽서를 동봉하고 있음. 이와 함께 우편물의 주소지가 지속적으로 불명확한 경우 대상자가 거주하고 있던 지방자치단체에 확인하고 있음.

○ 조직 및 인력

- JHPS/KHPS는 게이오대학 경제연구소의 패널데이터 연구센터에서 관리 운영하고 있음.
- 패널데이터 설계·해설센터는 제3세대 패널데이터의 설계· 실시·해석·공개를 통해
 - ① 가계소득변화나 계층 간 이동, 취업이나 고용·투자 등 행동 변화에 대한 실태 파악 및 국제비교
 - ② 경제이론으로 도출된 동학적 이론 가설의 검증
 - ③ 사회보장제도의 개정이나 법률·정책 변경 후 시간에 따른 정책평가 분석을 가능하게 하여 일본의 실증연구 수준을 비약적으로 향상시키기 위해 노력하고 있음.
 - 구체적으로 일본 전국을 대상으로 한 가계패널데이터와 상장기업의 기존 재무제표나 신규 개업기업에 대한 조사를 사용하여 기업에 관한 패널데이터를 구축하고 있음.
 - 또한 연구 결과는 콘퍼런스나 심포지엄을 통해 적극적으로 정부 발신을 하고 있음. 나아가 경제협력개발기구(OECD)나 룩셈부르크소득연구소(Luxembourg Income Study) 등 국제기관과의 공동연구를 진행해 일본을 대표하는 패널데이터에 근거한 연구거점으로서의 역할을 하고 있음.
- 센터 조직은 크게 데이터 구축 및 연계팀과 연구팀, 사무국으로 구성되어 있음.
 - 구체적으로 데이터 구축 및 연계반은 가계패널반, 연계반, 장수반으로 구성. 연구반은 워크라이프밸런스반, 자산격차반, 차세대육성반, 사회보장반, 재정반으로 구성되어 있음.

- 연구인력은 경제학 및 그 외 분야의 다양한 연구자를 참여시켜 패널데이터를 얻기 위한 조사표를 설계·분석하고 있음.

[그림 5-5] 게이오대학 패널데이터 연구센터 조직도



자료: 게이오대학 패널데이터 연구센터 홈페이지(조직도와 인력). <https://www.pdrc.keio.ac.jp/en/about/staffs/>에서 2018. 5. 15. 인출

□ 재원

○ KHPS

- 재원은 문부과학성·21세기COE(2003~2007년), 문부과학성·글로벌COE(2008~2012년)의 보조금과 과학연구비조성사업의 보조금이 이용되었음.

○ JHPS

- 문부과학성의 ‘특색 있는 공동연구거점 제 정비 추진사업’에서 연구거점으로 인정되어 위탁사업비가 지급되었음.
- 패널거점의 운영비용에는 크게 3개의 비용 항목이 있음. 각각 연구 활동비(조사비용, 연구원 인건비, 기타 설비비품이나 인쇄비용 등), 거점유지운영관리비(사무원 인건비, 임대료, 광열비 등), 법인관리비임.
- 이 비용은 연간 약 9,800만 엔 정도(초년도 조사비용으로 산출)이고, 연구 활동비가 전체의 약 75%, 거점유지운영비가 전체의 약 10%를 차지하였고, 특히 연구 활동비에 포함된 조사비용(초년도의 경우, 차년도 이후 약 8% 감액)이 전체의 약 50% 정도를 차지하였음.

○ JHPS/KHPS

- 통합 이후 과학연구비조성사업, 후생노동과학연구비보조금 및 정책과학종합연구사업, 과학기술진흥기구 및 사회기술연구개발센터 등의 연구조사 사업 등을 통해 보조금을 받고 있음.
- 과학연구비조성사업(특별추진연구) “장수사회의 세대 간 이 전과 경제격차: 패널데이터에 의한 정책평가분석” 2017~2021
- 과학연구비조성사업(특별추진연구) “경제격차의 다이너미즘: 고용·교육·건강과 재분배 정책의 패널분석” 2012~2016
- 과제설정에 의한 선도적 인문·사회과학 연구추진 사업·글로벌 발전프로그램 “국제비교 가능한 데이터에 의한 남녀공동 참가와 가족의 역할변화의 다원적 동학분석” 2013~2016.

- 후생노동과학연구비보조금·정책과학종합연구사업(정책과학추진연구사업) “취업 상태 변화와 적극적 노동시장정책에 관한 연구” 2014~2016
- 과학연구비조성사업(기반연구) “경제격차와 교육격차의 장기적 인과관계의 해명” 2016~2022
- “특색 있는 공동연구거점의 정비 추진사업(패널조사공동연구거점)기능 강화지원” 2016~2018
- 과학기술진흥기구·사회기술연구개발센터(사람과 정보의 에코시스템·프로젝트 기획조사) “사람과 AI시스템의 협동 태스크 모델(task model) 구축을 위한 조사” 2017

제4절 조사표의 주요 내용

□ 조사 내용

- KHPS는 주로 취업, 소비, 소득, 주택 등을, JHPS는 취업, 소득, 교육, 건강·의료 등의 내용을 조사해 옴.
- 2014년 이후 조사 항목을 통일하여 가족 구성, 개인 속성, 학력, 취업·취학 상태, 생활시간 배분, 부모와의 동거관계 등 포괄적인 주제로 조사함.
 - 일반 사항: 개인 속성, 가족 구성, 동거 여부, 중학생 이하 자녀가 있는 경우 소유하고 있는 책의 권수, 자녀와의 저녁식사 빈도, 부모 생존 여부, 부모 경제원조 여부, 부모로부터의 경제원조 여부 등

- 교육: 취학 여부, 유학 또는 이주 경험, 영어 수준, 장학금 경험 및 종류
 - 보건 및 기타: 흡연량, 음주량, 음주 습관, 운동 습관, 운동량, 암검진 여부 등
 - 복지 및 노동: 경제활동 여부, 취업형태, 근무시간, 근무 형태, 유급휴가일, 노동조합 가입 여부, 공적연금(국민연금, 고용보험, 개호보험, 국민건강보험), 개호(돌봄)가 필요한 가족에 대한 질문 등
 - 소득 및 지출: 예적금 액수, 유가증권 액수, 차입금 여부 및 이유, 자산 운용 현황, 증권시장에 관한 질문, 대출상환금액, 기부금
 - 예적금이란 우체국의 정책·정기·적립저금, 통상저금, 은행, 신용금고 등의 정기에예금·정기적금, 보통저금, 사내예금, 투자구좌, 저축구좌, 펀드 등을 말함.
 - 유가증권이란 주식(시가), 채권(액면), 주식투자신탁(시가), 공사채투자신탁(시가), 대부신탁·금전신탁(액면) 등을 말함.
 - 차입금: 차입금이란 금융기관 외의 곳, 친척, 친구 등에게 빌린 돈을 말함. 다만 할부금의 미납금이나 현금서비스에 의한 미납금은 제외.
 - 생활시간: 통근·통학, 가사, 육아, 자원봉사활동, 수면 시간 등
 - 생활만족도: 소득, 업무, 주거, 여가시간, 건강, 생활 전반, 자유와 평등
 - 재산: 저금액, 차입금액, 금융자산, 수입, 자산, 지출, 주거, 대출 등
- 조사 시작 이후 가족에 대한 정의 없이 조사한 결과, 경제적으로 독립해서 생계를 영위하는 자녀를 가족으로 인식하지 않고 조사

표에 정보를 기입하지 않는 케이스가 많이 발생함.

- 2017년부터 가족을 ‘부모, 배우자, 배우자의 부모, 형제자매, 배우자의 형제자매, 형제자매의 배우자, 자녀, 자녀의 배우자, 손자, 손자의 배우자, 이 외에 동거 중인 사람’으로 정의.

〈표 5-4〉 한국과 일본의 패널조사 내용

영역	세부 영역 및 세부 항목	
	한국	일본
가구 일반 사항	가구원 수	가구원 수
	가구원 일반 사항 - 가구원 번호, 이름, 가구주와의 관계, 성별, 생년·월, 교육 수준, 장애 종류 및 등급, 만성질환, 혼인 상태, 종교, 동거 여부	대상자 일반 사항 - 배우자 유무, 성별, 생년월일, 가구원 일반 사항 - 가구원 성별, 생년월일, 동거 여부, 취학·취업 상황, 가구주와의 관계
	가구 일반 사항 - 가구 형태(단독가구, 모부자가정, 소년소녀가장, 기타) - 기초보장수급 형태	가구 일반 사항 - 가구 형태 - 세대 변동 상황
건강 및 의료	건강 상태	건강 상태
	의료기관 이용 및 주요 병명 - 외래진료 횟수, 입원 횟수, 입원 일수, 병원에 입원한 주요 이유, 주로 이용하는 의료기관의 형태, 건강검진 횟수, 주요 병명	- 통원·입원 경험, 병원 이용 이유, 건강검진 횟수, 검진 결과, 체중 - 고액의료비 신청, 의료비공제 신청 - 정기적인 운동, 종류, 한 달 동안의 운동 일수, 하루 운동시간, 운동 장소
경제활동 상태	근로능력 정도	-
	경제활동 참여 상태	전월 경제활동 상황, 1년 전 경제활동 상황
	근로 A 취업자 일자리 특성 - 고용 형태, 근로시간 형태, 업종, 직종, 사업장 규모 근로 B	- 경영조직, 고용 형태, 근로시간 형태, 업종, 직종, 직위, 업무 내용, 사업장 규모, 사업장 소재지, 비정규직인 경우 그 이유, 계약 기간 유무, 유급휴가 일수, 노동

영역	세부 영역 및 세부 항목	
	한국	일본
	취업자 - 실직 경험 여부, 실직 사유, 현직 장 근속 연수, 근로 개월 수, 일한 달 평균 근로 일수, 주당 평균 근로시간	조합 가입 여부, 수입액, 급여 지급 형태(월급, 주급, 일급, 시급, 연봉), 상여금액, 주당 평균 근로 시간 - 부업 유무
	비취업자 - 비경제활동 사유	비취업자 - 비경제활동 사유
사회보험 가입, 퇴직금 적용	공적연금 가입	공적연금 - 연금지불액
	산재보험, 고용보험 가입	고용보험 가입
	퇴직(연)금 가입	개인보험 가입 여부 및 종류
	건강보험 가입	건강보험 가입 개호보험 가입 단계 개호가 필요한 가족 여부
주거	주택 유형 및 주거 위치	주택 유형 및 주거 위치 - 주택 유형, 정원 유무 및 면적, 건축 연수, 역세권, 배리어프리, 지대, 주택 구입 시기
	주택 구입비(보증금) - 주택(보증금) 가격, 월세액, 주택 구입비(보증금) 마련 방법, 대출상환액 연체 횟수	주택 구입비(보증금) - 주택(보증금) 가격, 주택(보증금) 마련 방법, 주택 구입 시기, 월세액, 계약기간 - 대출 여부, 대출상환총액
	주택 구조·성능·환경, 주거시설 - 주택 재질, 방음, 난방, 오염, 자연재해로부터의 안전 여부 등 - 방 개수, 주거 면적, 상하수도·부엌·화장실 등 주거시설의 종류 및 사용 형태	주택 구조·성능·환경, 주거시설 - 주거면적, 건물 층수 - 배리어프리 시설
생활비	비목별 지출 - 식료품비, 주거비, 광열수도비, 가구가사용품비, 피복신발비, 보건의료비, 교육비, 교양오락비 등	비목별 지출 - 식료품비, 주거비, 광열수도비, 가구가사용품비, 피복신발비, 보건의료비, 교통통신비, 교육비, 교양·오락, 용돈, 사적 이전

영역	세부 영역 및 세부 항목	
	한국	일본
		지출, 기타
	세금 및 사회보장 분담금	-
	총생활비	-
소득	종사상 지위별 가구원 근로소득	연봉(세전, 세후 금액), 가구원의 수입(월급, 자영·사업 수익)
	원천별 소득(근로소득 제외)	- 보험급여액, 퇴직금액, 기타 재산의 매각금액, 유가증권 매각손익 액수 - 월세 수입, 이자·배당금, 용돈, 실업급여·육아휴직급여 액수, 기타 수입
부채 및 이자	부채 - 부채형태(금융기관 대출, 일반 사채, 카드빚, 전세보증금, 외상 및 미리 탄 켜돈, 기타)	부채 - 부채 총액
	이자	-
	부채의 용도	부채의 용도 - 부동산 구입, 교육비, 결혼자금, 여행, 사업자금, 병, 생활비 등
재산	부동산 - 소유부동산 형태, 소유부동산 총액, 점유부동산 형태, 점유부동산 총액	부동산 - 소유부동산 형태, 가격
	동산 - 보유 농기계 종류 및 가격, 사육 농축산물 종류 및 가격	-
	금융자산 - 소유금융자산 형태(예금, 적금, 주식·채권, 타기 전 부은 켜돈, 기타), 자산별 금액	금융자산 - 예금, 적금 금액, 유가증권 액수
	기타 재산	금융자산 운용, 세대소득·금융자산 상황 파악 여부, 리스크를 동반하는 금융자산 보유 여부, 구입 여부 유가증권 시장에 대한 인상

영역	세부 영역 및 세부 항목	
	한국	일본
생활 여건	박탈지표	-
노인 (만 65세 이상 노인이 있는 가구)	노인부양의 어려움과 노인복지서 비스 이용 현황	부모 - 부모 생사 여부 - 부모에게 하는 경제원조, 이유 - 부모에게 받는 경제원조, 이유
아동 (만 17세 이하 아동이 있는 가구)	아동 양육의 어려움과 아동복지서 비스 이용 현황	(중학생 이하 아동이 있는 경우) - 자녀가 평소 즐기 위해 사용 하는 것(컴퓨터, 인터넷, 미술 품, 악기) - 주중 아동과 함께 저녁 식사 하 는 횟수, 집에 소장하고 있는 책 권수
	아동 건강	-
	사교육과 보육기관 이용 실태	-
가족	가족의 갈등	-
생활실태 만족 및 의식	생활실태 - 인터넷사용 여부, 유해환경 근 로경험 여부	생활실태 - 인터넷 사용 여부, 사용수단
	생활만족 - 건강 상태 만족, 가구소득 만 족, 주거환경 만족, 가족 관계 만족, 직업 만족, 사회적 친분 관계, 여가생활, 전반적인 생활 - 경제적 생활 상태(조사 기준 시 점, 5년 후), 복지 인식	생활만족도 - 세대소득, 직업, 주거, 여가시간, 건강, 생활 전반, 자유와 평등 - 행복감(10점 척도), 삶의 보람, 삶의 희망(5점 척도)
사회적 환경에 대한 인식	사회적 자본 - 신뢰, 지역사회 이웃에 대한 도 움 의사, 비선호시설 수용 정도 - 기부 혹은 자원봉사활동 여부, 연간 기부액, 자원봉사활동 연 간 횟수 - 도움을 받은 곳, 도움을 준 사 람 수, 도움 종류와 정도	- 자원봉사활동 여부, 자원봉사활 동 연간 횟수, 기부활동 여부
	부모님과 의 접촉 정도 - 비동거 부모님 유무, 비동거 부 모님과 의 왕래 횟수	-
	생활시간	생활시간

영역	세부 영역 및 세부 항목	
	한국	일본
	- 취업활동, 집안일, 자녀돌보기, 가족돌보기, 교육과 훈련, 여가 활동	- 통근·통학 평균시간, 가사시간, 육아시간, 업무를 위한 연수·학습시간, 자원봉사활동, 평일 수면시간, 휴일 수면시간
	성 역할	-
생활 습관, 가족관계 및 정신건강	생활습관 - 흡연, 음주, 음주로 인한 문제	흡연 - 흡연 여부, 타르양, 니코틴양, 가격, 하루 흡연량 음주 - 음주 습관
	정신 건강 - 우울감, 자아존중감	-
	부부관계 및 가족관계 - 부부 폭력 경험, 부부 폭력 피해 경험	-
	자살 생각, 계획, 시도	-
교육	교육 - 최종 학력, 출신 고등학교의 유형과 소재지, 출신 대학의 전공 계열과 소재지	교육 - 최종 학력, 출신학교 소재지 - 유학 경험, 자신의 영어 수준 - 장학금
기타		개인보험 가입 여부 - 화재보험(집, 재산), 지진보험, 생명보험 지진보수공사 실시 상황

자료: 한국보건사회연구원·사회복지연구소(2018). 한국복지패널 12차년도 조사자료 유저가이드, 게이오대학 패널데이터 연구센터(2016). 일본가계패널조사(JHPS/KHPS) 조사표를 바탕으로 재기술함.

제5절 패널 활용

□ 홈페이지 운영

- 패널 데이터 자료 이용 및 연구 결과, 연구 성과 등을 홈페이지에 올려 다운로드하여 이용할 수 있음.
 - <https://www.pdrc.keio.ac.jp/>
 - 패널데이터는 게이오대학 패널데이터 연구센터 홈페이지에서 신청하고 다운로드하여 이용할 수 있음.
 - 또한 일본가계패널데이터 외에 센터에서 관리하고 있는 ‘일본 아동패널조사(JCPS)’, ‘동일본대지진에 관한 특별조사(GEES)’, ‘신규개업패널조사’ 등의 데이터도 이용 가능함.

□ 패널데이터 제공 방식

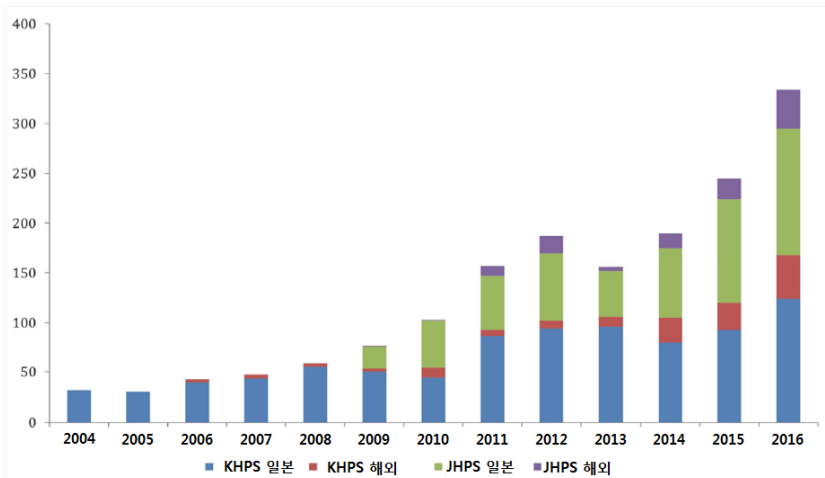
- 게이오대학 패널데이터 연구센터는 비영리·학술 목적의 이용에 한해 국내외의 대학 또는 국공립·민간 연구기관에 소속되어 있는 연구자·대학원생·학부생(연구생·연구원 등 포함)에게 데이터를 제공함.
- 이용에 있어서는 계약서에 기재되어 있는 사항을 의무로 하고, 계약사항을 위반한 경우 데이터 이용을 금지하고 있음.
- 홈페이지에서 신청한 후, 센터의 심사를 거쳐 이용 승인이 난 경우에 한하여 신청한 데이터를 다운로드하여 사용할 수 있음.
 - 신청이 승인된 날로부터 1년까지 다운로드하여 사용할 수 있음.
 - 이용승인 종료 시에는 일본가계패널 홈페이지에 제공된 데이터를 이용한 분석 결과를 등록해야 함. 등록하지 않을 경우 이

후 데이터 신청이 불가함.

- JHPS/KHPS는 지금까지 많은 연구자에게 이용되어 사회과학의 다양한 분야에서 성과가 축적되고 있음. 조사를 시작한 최초 수년간은 주로 일본 국내 이용자에게 이용되었지만, 최근에는 해외저널이나 국내외 학술기관과의 연계·교류활동 등으로 일본의 대표적인 가계패널조사로서 인지도가 높아져 데이터 이용 건수가 증가하고 있음.
- 2016년도에는 총 334건의 데이터 이용 건수가 있었음. 이용 건수는 데이터를 이용해 작성된 논문의 건수임.

[그림 5-6] JHPS/KHPS 이용 실적

(단위: 데이터 제공 건수)



자료: 일본가계패널조사 뉴스 제14호. https://www.pdrc.keio.ac.jp/uploads/JHPSKHPS_panelnews_No.14.pdf에서 2018. 5. 15. 인출

□ 외부 패널과의 협력

- 국제연계의 일환으로 LIS의 데이터베이스와 연계.
- CNEF와 연계하여 국가 간 비교가 가능한 패널이 되는 데 포함될 예정.
- 경제협력개발기구(OECD)의 소득변동 국제비교연구 등에서도 일본을 대표하는 데이터로 활용되고 있음.
 - 게이오대학 패널데이터 연구센터의 외부 연계기관
 - 일본정책금융공단종합연구소
 - 국립교육정책연구소
 - 히토쓰바시대학 경제연구소
 - 교토대학경제연구소
 - 도쿄대학 사회과학연구소 사회조사·데이터아카이브 연구센터
 - LIS
 - Centre for Educational Research and Innovation (CERI) , OECD
 - Directorate for Employment, Labour and Social Affairs(ELSA), OECD
 - Cross-National Equivalent File(CNEF), Ohio State University
 - International Inequality Institute, LSE

□ 세미나 및 워크숍 개최

- 게이오대학 패널데이터 연구센터에서는 정기적으로 협력기관과

의 세미나 및 워크숍을 개최하고 있음.

○ 2017년 패널조사를 수행하는 연구기관의 국제협력과 발전을 위해 한국보건사회연구원과의 공동세미나를 실시해 JHPS/KHPS의 개요를 설명하고 이를 이용한 연구 결과를 발표함.

- 본 세미나에서 발표된 내용은 다음과 같음.

〈표 5-5〉 International Seminar on KOWEPS & JHPS/KHPS

내용	발표자
1세션. 일본가계패널연구(JHPS/KHPS)	
일본 패널연구 소개	Myoung-Jung KIM (NLI Research Institute)
JHPS/KHPS의 데이터 구조와 근거 기반 정책	Fang HE (Keio University)
JHPS/KHPS를 활용한 빈곤과 불평등 분석	Kayoko ISHII (Keio University)
2세션. 한국복지패널(KOWEPS)	
한국복지패널 소개	김태완, 박형준 (한국보건사회연구원)
한국복지패널을 활용한 중단분석 연구 사례	김문길 (한국보건사회연구원)
한국과 일본의 불안정 노동시장 비교 연구	백승호 (가톨릭대학교 사회복지학과)
3세션. 종합토론	
손창균(동국대학교 응용통계학과) 유태균(숭실대학교 사회복지학과) 조준용(한림대학교 사회복지학과) 여유진(한국보건사회연구원)	

자료: 한국보건사회연구원(2017), KOWEPS&JHPS/KHPS 공동국제학술회의 자료집

○ 해외 교수들을 초청하여 세미나 및 워크숍을 진행.

○ 2018년도 개최 및 개최 예정인 세미나 및 워크숍은 다음과 같음.

〈표 5-6〉 게이오대학 패널데이터 연구센터의 세미나·워크숍 일정(2018년도)

일정	내용
2018. 8. 1.(수)	유아교육연구 문제에 대한 소규모 컨퍼런스
2018. 7. 31.(화)	‘유아기교육의 경제학’
2018. 7. 1.(일)	일본학술진흥회 ‘인문학·사회과학 데이터 구조(infrastructure) 구축 프로그램 키포프 심포지엄 데이터 기반 정비’
2018. 5. 24.(목)	불균형한 노화? 인도의 부(富) 불평등에 대한 첫 번째 견해
2018. 4. 25.(수)	WIC(Women, Infants, and Children) 참여와 가구 식품 구매의 상대적 품질
2018. 3. 3.(토)	노동정책과 실증분석
2018. 3. 2.(금)	불평등의 경제적 사회적 결과
2018. 2. 28.(수)	OECD 국가에서 불평등이 증가하는 원인과 기제
2018. 1. 12.(금)	아동과 가족의 평등 문제에 대한 소규모 컨퍼런스: 중국, 한국, 일본 그리고 세계로부터의 근거

자료: 게이오대학 패널데이터 연구센터 홈페이지(이벤트) <https://www.pdrc.keio.ac.jp/en/events/>에서 2018. 5. 15. 인출

□ 활용 실적

- 일본가계패널조사는 국내외 연구 등에 활용되고 있으며, 일본 국회에서는 예산위원회, 소득격차 관련 자료 등에 활용되고 있음.
- 최근 해외 연구기관으로부터의 데이터 액세스가 증가하여 경제협력개발기구(OECD) 소득변동 국제비교연구 등에서도 일본을 대표하는 데이터로 활용되고 있음.
 - OECD(2008) “Growing unequal? Income Distribution and Poverty in OECD Countries”
 - 소득동태의 국제비교분석
 - OECD(2015) “Employment Outlook 2015”
 - 근로소득의 변동
 - OECD(2015) “In It Together-Why Less Inequality

Benefits All.”

- 비정규고용과 소득격차의 국제비교분석

○ 매년 조사 결과물을 발간하고 있으며 보고서, 설명회, 심포지엄 등을 연 1~2회 정도 진행, 홈페이지에서 소식 및 결과 제공.

제6절 복지패널에 주는 시사점

□ 국제적인 패널로서 도약

○ 한국복지패널이 국제적으로 활용되기 위해서는 외부 패널과의 협력이 필요.

- 한국복지패널을 사용한 국제비교 연구 등 다양한 주제 발표가 필요.

○ 국제기관과의 데이터 공유와 분석 결과 제공.

○ 해외 패널 관련 연구소와 연계한 세미나 및 워크숍 개최.

○ 일본은 해외에서의 패널 활용을 파악하기 위해 외부에 패널 제공 시 활용 논문에 대한 보고를 하도록 함. 한국복지패널도 외부에서 얼마나 어떻게 활용되고 있는지에 대한 파악과 공개를 통해 복지패널의 위상을 알림과 동시에 활용도를 높이기 위한 방안을 마련해야 할 것임.

제 6 장

대만 Panel Study of Family Dynamics(PSFD)

제1절 조사 배경

제2절 표본 및 가중치

제3절 조사 방식 및 관리

제4절 조사표의 주요 내용

제5절 패널 활용

제6절 복지패널에 주는 시사점

6

대만 Panel Study of Family Dynamics(PSFD) <<

제1절 조사 배경²⁶⁾

□ 조사 설계 배경

○ PSFD는 중국 가정의 성인표본을 대상으로 하는 고정표본 추적조사(panel)로 청소년부터 노인까지를 대상 연령층에 포함함. 이 외에 본 계획은 주요 표본을 토대로 조사 대상을 자녀 등의 친족까지 확대하고 다시 추적 방문조사를 함.

- 본 장에서는 PSFD를 통해 중국인 가정의 경제, 사회, 심리, 관습 등에 대한 전반적인 내용을 살펴봄.²⁷⁾

○ 본 프로젝트의 구상은 중국인 사회의 가정 형태, 구조, 상호 패턴에 기초를 두고 있고, 이는 서양의 사회보다 훨씬 복잡함. 또한 그 안에 내포된 이론적인 패턴 역시 서양 사회의 그것에 기인한 것보다 복잡한 편임. 이에 중국인 가정의 동태 데이터베이스를 구축하고 현지 사회 환경과의 일치점을 찾으며, 나아가 서양의 주류 학술가치로서도 수긍 가능한 연구 방향을 정립할 수 있기를 희망함.

26) 중국 가정 동태 데이터베이스(PSFD)의 조사 설계 배경, 특징, 조사 방법과 현황에 대한 내용은 PSFD 홈페이지 중 조사 배경(<https://psfd.sinica.edu.tw/web/>) 부분과(2018. 7. 31. 인출), 우약용·장영화(于若蓉·章英華, 2014), ‘화인가정동능자료고: 대만방문계화(華人家庭動態資料庫: 台灣訪問計畫)’의 자료를 참조하였음.

27) 중국 가정 동태 데이터베이스(PSFD)는 대만에서 시작된 패널조사가 향후 중국 대륙의 성(상하이, 절강성, 푸젠성)으로 확대된 것임. 이러한 부분이 대만인을 대상으로 한 조사인지 중국인을 대상으로 한 조사인지 혼란을 줄 수 있으나, 영토적으로 대만과 3성에 거주하는 가족(family)을 대상으로 한 패널조사라는 점에 유념할 필요가 있음.

- 대만 자료 수집을 시작으로 가정을 기본 단위로 설정하고 시준별 추적조사를 진행하며 대만 가정의 동태 데이터베이스를 구축한 후 향후 기타 중국 사회로 (범위를) 확대할 것임. 자료 수집의 경우 한편으로는 기존의 서양 유관 이론이 대만 사회에 적용 가능한지 살펴보고 다른 한 편으로는 데이터가 나타내는 사실을 기반으로 하여 서양의 이론과 다른 새로운 이론을 발굴하고 추출하여 구축하는 데 목적이 있음.
- 1998년, 중앙연구원의 주도하에 심리, 사회, 경제 등의 분야에서 다수의 학자들이 본 프로젝트에 참여하기 위해 집결하였음.
- 본 데이터베이스가 본토의 특색과 국제적 주류의 추세와 연계될 수 있도록 해외의 지명도 있는 학자들을 초빙하여 고문 역할을 할 수 있도록 하였고, 프로젝트의 설문 설계 등에 대해 감수를 받음.

□ PSFD의 특징

- 프로젝트 참여자 모두가 ‘공공재(公共財)’라는 이념을 설정하여, 데이터의 오류 여부 검사 완료 후 즉시 공개하고 국내외 학계의 인사들에게 자료를 활용할 수 있도록 함(于若蓉·章英華, 2014, p.2).
- 매 조사는 인터뷰 완료 시점부터 데이터 해석 완료까지 약 1년 정도의 시간이 소요됨. 이는 비슷한 유형의 국내외 대형 데이터조사와 비교하였을 때 상당히 빠른 속도로 진행된 것임.
- 현재까지 대만에서 진행된 패널조사 중 기간이 가장 길고, 방문 차수가 가장 많음.
- 가정 중심: 가족 구성원 중 한 명이 대표로 자신, 배우자, 부모, 자

녀, 형제자매 등 가족관계에 대해 응답.

- 주요 내용은 교육, 직업, 가정수입, 가족의 상호작용, 생활 능력 등

○ 부모, 자녀 표본 포함: 조사 대상은 표본추출 시 주 표본 외에 16세 이상의 자녀도 포함.

○ 대만 외에 중국대륙(동남중국) 자료 포함: 대만 조사 외에 동시에 푸젠성, 상하이, 절강성(동남중국)에서 추적방문조사 진행.

□ 조사현황 및 미래전망

○ 대만 중국인 가족 조사

- 대만에서 진행하고 있는 방문 인터뷰 프로젝트는 1953년부터 1964년까지 출생한 총 1,000개의 샘플에 대해 방문조사를 진행하는 것으로 1999년 마무리되었고, 2000년에는 1935년부터 1954년까지 출생한 1,959개의 샘플에 대한 방문조사가 마무리되었음. 2년간 완성된 표본은 대만의 주 표본으로서 이후 매년 이루어지는 추적조사의 중요한 기초이기도 함.
- 두 세대의 관계가 가정 관계에서 중요한 일환이라는 점, 그리고 두 세대의 관계가 가정의 생명주기에 따라 변한다는 사실에 비추어 2000년 7월부터 8월 사이 앞서 서술한 두 종류의 표본인 16세에서 22세 사이의 자녀에 대해 방문 인터뷰를 하고, 격년 조사 방식으로 자녀 표본에 대한 추적조사를 하였음.
- 2001년부터는 주요 표본의 범위에 16세의 자녀 역시 자녀 표본으로 포함시켜 부모와 자녀의 누적 대조자료로 활용하였음.
- 2001년부터 2005년 사이에는 앞서 설명한 표본의 추적방문

조사 외에도 2003년에 1964년부터 1976년 사이에 태어난 표본을 주 표본으로 새롭게 편입시켜 첫 번째 방문 조사를 함.

- 2004년에는 자녀 표본 중 25세(1978~1979년 출생자)를 주 표본군에 편입시켜 설문에 적용함. 여기에서 주요한 고려 사항은 25세에 달한 표본은 이미 독립적인 경제정책, 인간관계를 스스로 설계할 수 있어 주 표본에 포함될 수 있는 것으로 판단하였다는 것임. 또한 2005년 5월에 프로젝트 보충조사를 하였는데 표본 유실률을 줄일 수 있었기에 데이터의 사용가치가 더욱 상승하였음.
- 2011년까지 줄곧 3세대 연령군의 주 표본을 기준으로 제13, 12 및 제9차 방문을 하였음. 점차 증가하는 유실 표본을 보충하기 위해 2009년에 새롭게 표본을 추출하는 방식으로 표본 수를 늘렸고(이 표본의 경우 1977년부터 1983년까지 태어난 사람을 대상으로 함) 이후 주 표본의 추적 조사에 포함시킴. 2011년 해당 표본에 대해 제3차 방문을 진행하였음.
- 추적조사에서 얻은 데이터는 여러 해 누적된 시간이 지나야 비로소 성과가 나타남. 프로젝트는 지금(2014년)까지 이미 15년간 실행되었으며, 조사 자료와 교류자료의 사용경험을 널리 알리기 위해 2005년 12월 2일과 3일 이틀간 ‘중국인 가정 동태 데이터베이스 데이터 사용 워크숍’을 개최함. 회의에서는 대만과 대륙의 인터뷰 계획을 소개했을 뿐만 아니라 학계 관계자들이 데이터베이스의 특성과 분석 방법을 일목요연하게 이해할 수 있도록 하였고 이와 동시에 본 데이터베이스를 활용하여 완성된 연구논문도 발표되었음.

○ 대륙 중국(동남중국)인 가정 조사

- 2003년 본 프로젝트는 중국 사회과학원의 인구 및 노동경제 연구소와 협력하여 상하이, 푸젠성 및 절강성 등 3성(시) 및 대만 RI2003 조사표를 설계의 바탕으로 삼아 양안(대만과 대륙) 국민 정서의 차이점을 염두에 두고 수정한 조사표를 이용해 방문 조사를 진행하였음.
 - 조사는 2003년 기준 만 25~68세의 성인을 선발하고 5,000건의 표본을 추출하여 방문 조사를 진행할 예정이었으나 사스(SARS)의 영향으로 조사가 연기되었고 2004년 하반기 완료됨.
 - 3성에서 각각 완료된 표본의 수는 아래와 같음.
 - 상하이: 903개
 - 절강성: 1,856개
 - 푸젠성: 1,925개
- PSFD 구축은 1999년 및 2000년 국가과학위원회 인문처, 센터와 장징궈(蔣經國)기금회에서 공동으로 투자하고, 중앙연구원경제연구소를 집행기관으로 하여 프로젝트를 진행함. 현재는 중앙연구원 인문사회학 연구센터 이슈조사센터하의 ‘중국인 가정 연구 프로그램’의 일환으로 추진되고 있음(于若蓉·章英華, 2014, p.2).
- 이러한 프로젝트를 대만 가정에 대한 설문조사의 기점으로 설정하고 그 범위를 대만 외 중국인 가정으로 확대하여, 구축된 데이터를 국내외 학자들이 사용하기를 기대하고 있음.
- 중국인은 전 세계 인구의 4분의 1을 차지하고 있고 가정환경 또한 독특한 특성을 가지고 있으므로 다양한 분석이 시도될 수 있음.

- 1999년 조사가 시작된 이래, 13년의 방문 인터뷰 계획이 수립되었고 상당량의 정보가 축적되었음.

□ 현지조사 방법

- PSFD의 현지조사는 중앙연구원 조사연구센터의 조사팀과 PSFD 설계팀이 협력하여 진행함. 중앙연구원 조사연구센터에서는 조사 관리자와 조사원 등의 인력과 인력 연수를 지원함. 그 중 조사 관리자는 다년간의 경험과 연구센터와 장기 협력관계에 있는 전문가로 구성하였고, 조사원은 매회 조사 전 대외모집을 통해 모집함. 인력 연수과정에서는 실무훈련과 조사표에 대한 이해를 중심으로 교육을 실시함.
- 현지조사는 매년 1월 말에서 2월 초에 시작하여 약 1개월간 진행함. 2004년, 2005년에는 완성률을 높이기 위해 1월 및 2월 정식 방문을 마친 후, 미완성 샘플의 경우 4월 및 5월에 보충 조사를 함. 매번 방문조사를 마친 후 ‘100분의 100 반복 조사 방식’으로 피조사자에게 전화를 걸어 실제로 방문이 이루어졌는지 확인하는 절차를 거침.

□ 조사 대가

- 조사 완료율을 높이기 위해 조사원은 50~150 TWD 상당의 답례품을 증정하거나 추첨행사, 연하카드, 조사 결과에 따른 보고서 등을 우편으로 전송함(于若蓉·章英華, 2014, p.13).
- 관리감독을 위해 조사원은 조사 대가에 대한 모든 부분을 보고하도록 함.

□ 데이터 분석 결과

- 현재 대만 지역의 데이터는 이미 분석이 완료되어 1999년부터 2011년까지의 설문조사 데이터를 공개하고 있음. 2012년부터 2014년까지의 설문조사 데이터는 현재 검정 단계에 있으며, 정오(正誤) 검사 완료 후 즉시 공개할 예정임.
- 데이터 검정 작업은 본 프로젝트(<http://psfd.sinica.edu.tw/>)와 인문사회학 연구센터(Academia Sinica Research Center) ‘학술조사데이터베이스’의 “Academic Survey Database”(<http://srda.sinica.edu.tw/>)의 협력으로 진행되고 있으며, 후자를 통해 자료를 해석하여 공개함. 공개된 파일은 조사표, 코드북, 데이터 파일(SAS, STATA, SPSS 시스템 파일 포함) 등임.
- 대륙지역 조사 자료의 경우, 2004년 첫 조사 데이터 및 2006년 추적조사 데이터는 ‘학술 조사 연구 데이터베이스’에서 분석 결과를 공개함. 2011년 설문조사 데이터는 이미 검증이 완료된바, 이 정보에 관심이 있는 연구자는 프로젝트팀과 협의하여 데이터를 얻을 수 있음. 2013년 조사 데이터는 양측이 현재 협력하여 데이터를 정리하는 과정에 있음.
- 개별 설문 인터뷰의 개별적인 응답자료 외에도 1999년부터 축적된 자료는 사용자의 편의를 위해 분석되었으며, 2004년에는 각각의 다른 조사에 대한 자료의 병합이 이루어짐. 데이터를 어떻게 병합하는가 하는 문제를 고려하는 경우에는 일반적인 연구자의 요구에 따라 프로젝트를 진행하였음.
- 일반 연구자가 사용하는 몇 가지 유형의 병합 데이터가 있음. 동일한 표본 그룹에 대한 데이터 추적, 여러 표본 그룹에 대한 동일

한 차수의 병합 데이터가 그것인데, 현재까지 병합되어 구성된 데이터는 총 4종이며 아래와 같음.

- 데이터세트1(RI-2000, RII-2001 및 RIII-2002의 병합 파일)
- 데이터세트2(RI-1999, RII-2000, RIII-2001 및 RIV-2002의 병합파일)
- 데이터세트3(RI-1999, RI-2000 및 RI-2003의 병합파일)
- 데이터세트4(CI-2000 및 CII-2002의 병합파일)

○ 위의 4종의 데이터에 대한 파일과 코드북, 자료파일(SAS, STATA, SPSS 시스템파일 포함) 등 연관 내용은 중영(中英) 버전으로 편집될 예정이며 그 내용은 학자들이 연구에 활용할 수 있도록 제공될 예정임.

○ 병합 파일을 사용하는 것 외에도 본 프로젝트는 데이터 검정 과정을 진행함으로써 데이터의 품질을 보증함. 그러나 병합 데이터 파일은 데이터 내용의 신뢰성으로 인한 의문의 여지가 여전히 존재하므로 현재 내부적인 검증이 진행되고 있음. 따라서 아직 정식으로 다운로드할 수 있도록 개방되지 않았으며, 필요한 경우 사용자는 협의 과정을 통해 프로젝트 어시스턴트에게 상담 가능함.

제2절 표본 및 가중치²⁸⁾

□ 대만지구 조사

○ 2종 표본

- 주 표본: 데이터베이스에서 추출.
- 자녀 표본: 아버지 혹은 어머니의 조사로 인해 조사를 받은 자녀.

○ 3종 조사

- 제1차 방문 조사: 조사 코드 RI 혹은 RCI, 조사표본을 사용하여 추적조사 데이터베이스에서 1차 조사. 주목적은 기초 정보 수집임.
- 추적 주요 조사: 조사 코드 RII, RIII, RVI...RR, 추적방문의 주요 조사 진행.
- 자녀 조사: 조사 코드 CI, CII, ..., CIX, 자녀표본(16~24세) 조사 진행, 주목적은 자녀의 학력 혹은 직업 현황 파악임.

○ 표본 구분

- 주 표본: '제1차 방문조사(RI)'를 사용하여 조사 진행, 차후에 '추적 주요 조사'를 사용하여 조사 진행.
- 자녀 표본: 16~24세 '자녀 표본'을 이용한 방문 조사, 25세는 '제1차 방문조사' 후 '추적 주요 조사' 진행.
- 추적 주요 조사의 데이터베이스는 주요 표본과 만 25세 자녀 표본으로 이루어져 있음.

28) 1999년부터 2016년까지 PSFD의 표본이 변화한 내용은 조사 내용 전반을 소개한 황력가(黃奕嘉, 2017). '화인가정동능자료고: 방문가구여문권설계(華人家庭動態資料庫: 訪問架構與問卷設計)'를 참조하였음.

○ 표본 현황 - 주 표본

- 5차 추출: 1999/2000/2003/2009/2016 추출.

〈표 6-1〉 차수별 표본 추출 방법

(단위: 명)

연도	출생 연령	예측 완성 수	실제 완성 수	표본 추출 방법
1999	1953 ~1964	1,000	999	- '분층3단등기율(分層3段等機率)' 표본을 원칙으로 합격샘 플을 추출하며, 이 외에 1.5~2배의 표본을 함께 준비. "대만지역발전유형연구"의 분할법을 원칙으로 분할하며, 1997년 말 정부 인구통계를 기준으로 표본을 추출.
2000	1935 ~1954	2,000	1,959	
2003	1964 ~1976	1,300	1,152	1999년의 촌리(村里) 표본에 의거하여 합격 표본을 취함. 단계별 완성률에 의거. 팽창비율법을 사용하여 표본 추출. 팽창 비율 1.6~2.1배 사이의 매 추출 사례는 필수 조사 대상.
2009	1977 ~1983	2,000	2,092	- '대만지역유형연구'의 지역 분할법과 2008년 6월 인구통 계서의 계산법을 기준으로 각 단계의 인구 비율을 계산. '향, 진'(중국 지역 구분), '촌, 리'의 표본을 추출하고 마 지막으로 팽창비율법을 이용하여 표본 사례를 추출, 팽 창비율은 1.5~2.5배 사이.
2016	1984 ~1991	2,000	1,973	

자료: 黃奕嘉(2017), p.6.

- 분층3단등기율(分層3段等機率)

- 1단: 향·진·시(구)(鄉·鎮·市)에서 추출.
- 2단: 촌리(村里)에서 추출.
- 3단: 출생연도가 부합하는 사람 추출.

〈표 6-2〉 표본 유실: 1999 표본(1953~1964)

(단위: 명, %)

조사 회차	방문 연도	성공 표본	유실 표본	유실률	누적 유실률
1	1999	999	-	-	-
2	2000	802	197	19.72	19.72
3	2001	726	76	7.56	26.97
4	2002	782	-56	-5.36	21.33
5	2003	751	31	3.12	24.45
6	2004	706	45	4.53	28.98
7	2005	677	29	2.92	31.89
8	2006	663	14	1.42	33.30
9	2007	637	26	2.62	35.92
10	2008	616	21	2.11	38.08
11	2009	594	22	2.21	40.24
12	2010	564	30	3.02	43.26
13	2011	554	10	1.01	44.27
14	2012	545	9	1.62	45.17
15	2014	518	27	4.95	47.89
16	2016	488	30	5.79	51.15

자료: 黃奕嘉(2017), p.7.

〈표 6-3〉 표본 유실: 2000 표본(1935~1954)

(단위: 명, %)

조사 회차	방문 연도	성공 표본	유실 표본	유실률	누적 유실률
1	2000	1,959	-	-	-
2	2001	1,730	229	11.69	11.69
3	2002	1,642	88	4.49	16.18
4	2003	1,566	76	3.38	20.06
5	2004	1,489	77	3.93	23.99
6	2005	1,396	93	4.75	28.74
7	2006	1,341	55	2.81	31.55
8	2007	1,288	53	2.70	34.25
9	2008	1,259	29	1.48	35.73
10	2009	1,202	57	2.91	38.64
11	2010	1,158	44	2.25	40.89
12	2011	1,144	14	0.71	41.60
13	2012	1,066	26	2.38	45.58
14	2014	977	89	8.35	50.13
15	2016	910	67	6.86	53.55

자료: 黃奕嘉(2017), p.8.

〈표 6-4〉 표본 유실: 2003 표본(1964~1976)

(단위: 명, %)

조사 회차	방문 연도	성공 표본	유실 표본	유실률	누적 유실률
1	2003	1,152	-	-	-
2	2004	832	320	27.78	27.78
3	2005	808	24	2.08	29.86
4	2006	768	40	3.47	33.34
5	2007	745	23	1.99	35.33
6	2008	716	29	2.52	37.85
7	2009	704	12	1.04	38.89
8	2010	685	19	1.65	40.54
9	2011	654	31	2.69	43.23
10	2012	652	2	0.31	43.40
11	2014	602	50	7.69	47.74
12	2016	562	40	6.64	51.22

자료: 黃奕嘉(2017), p.9.

〈표 6-5〉 표본 유실: 2009 표본(1977~1984)

(단위: 명, %)

조사 회차	방문 연도	성공 표본	유실 표본	유실률	누적 유실률
1	2009	2,182	-	-	-
2	2010	2,172	461	21.13	21.13
3	2011	1,685	36	1.65	22.78
4	2012	1,644	41	2.43	24.66
5	2014	1,539	105	6.39	29.47
6	2016	1,497	42	2.73	31.39

자료: 黃奕嘉(2017), p.10.

- 유실 원인: 방문 거절, 임시적 피조사자의 부재(재외 근무 등), 사망, 신체 혹은 정신적인 문제로 조사 불가, 이주 및 연락처 누락 등

〈표 6-6〉 표본 현황: 자녀표본

(단위: 명)

방문 연도	자녀 조사	제1차 방문	추적 주 조사	통계
2000	1,375	-	-	1,375
2002	1,378	-	-	1,378
2004	1,093	298	-	1,391
2005	-	167	320	487
2006	958	-	467	1,425
2007	-	322	462	784
2008	809	-	764	1,573
2009	-	269	746	1,015
2010	801	-	952	1,753
2011	-	217	903	1,120
2012	751	-	1,085	1,836
2014	650	342	1,035	2,027
2016	578	217	1,372	2,112

자료: 黃奕嘉(2017), p.11.

□ 동남중국(東南中國) 조사

○ 동남중국 조사 요약

- 조사지역: 상하이, 절강성, 푸젠성

○ 표본 추출

- 대륙 5차 인구조사자료를 기반으로 비농업인구 점유 비중으로 단계를 분할한 각 단계에서 확률비례추출 방법(Probability Proportionate Sampling: PPS)으로 표본 추출. 현, 구(중국 지방 구분 단위) 단위로 추출.
- 선택된 현, 구 단위는 PPS 방법으로 도로와 향, 전(중국 지방 구분 단위로 현, 구의 하위 단위임)을 추출.
- 다시 PPS법으로 향, 전에서 촌의회, 주민의회(중국 대륙 농촌 지역의 자치조직)를 추출.

- 추출된 촌·주민의회 중 등거리추출법을 사용하여 가구를 추출
- 키쉬그리드 추출법(Kish Grid Sampling)²⁹⁾으로 표본을 추출하여 합격한 가구에 대해 방문 조사를 진행함.
- 대상자의 사유로 방문이 거절될 경우 혹은 부재중으로 조사를 완성할 수 없는 경우 다시 표본을 추출하거나 가구를 변경하여 표본을 보충하지 않는 것을 원칙으로 함.

○ 표본 구조는 표본의 출생 연도의 범위는 대만 조사 대상의 범위와 동일함(1935~1976).

- 격년 1회 방문조사를 원칙으로 하였으나, 업무상의 원인으로 실제 방문은 2004, 2006, 2011, 2013, 2017년 총 5차로 진행.
- 중국사회과학연구원과 노동경제연구소의 협력으로 진행.
- 2004년 1차 방문 시 푸젠성 1,925건, 절강성 1,856건, 상하이 903건 완료.

〈표 6-7〉 동남중국 조사 표본의 규모와 유실률

(단위: 명, %)

조사 회차	방문 연도	성공 표본	유실 표본	유실률	누적 유실률
1	2004	4,864	-	-	-
2	2006	4,370	314	6.70	6.70
3	2011	3,871	499	11.42	17.36
4	2013	3,627	224	5.79	22.57

자료: 黃奕嘉(2017), p.14.

29) 사전에 할당된 임의의 숫자표를 사용하여 조사 대상자를 선정하는 방법임.

제3절 조사 방식 및 관리

□ 조사 방식: 조사원을 통한 방문 인터뷰

□ 조사 주기([그림 6-1] 참조)

[그림 6-1] 조사 주기와 조사 대상

조사 대상 나이(세)	조사 진행 연도															
	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2014	2016
24~43		RI- 2000	RII- 2001	RIII- 2002	RIV 2003, RV 2003	RVI- 2004, RV- 2004	RR- 2005	RR- 2006	RR- 2007	RR- 2008	RR- 2009	RR- 2010	RR 2011	RR 2012	RR 2014	RR 2016
42~53	RI- 1999	RII- 2000	RIII- 2001	RIV 2002	RI- 2003 (新抽樣)	RCI- 2004	RCI- 2005		RCI- 2007		RCI- 2009		RCI 2011		RCI 2014	RI 2016 (新抽樣)
53~65																
65																
66																
67~68																
69																
70																
71																
72																
73																
74																
75																
76																
77~78																
79~80																
81~82																
83~84																
85~86																
87~88																

주: 파랑(주 샘플), 노랑(자녀 샘플), 초록(신규 샘플), 25세 이상의 자녀 샘플을 대상으로 처음 주 샘플조사를 진행함.
자료: PSFD 홈페이지(조사 구조) <https://psfd.sinica.edu.tw/web/check.htm>에서 2018. 7. 31. 인출

□ 조직 및 인력

〈표 6-8〉 PSFD 연구 조직과 인력

	이름	직책	연락처
1	于若蓉 Yu Ruo Rong	교수	yurr@gate.sinica.edu.tw
2	簡錦漢 Jian Jin Han	교수	
3	章英華 Chang Ying hua	교수	ethwa@gate.sinica.edu.tw
4	蔡明璋 Cai Ming Zhang	교수	
5	周玉慧 Zhou Yu Hui	교수	jade@gate.sinica.edu.tw
6	張峰彬 Zhang Feng Bin	교수	
7	陶宏麟 Tao Hong Lin	교수	
8	胡立中 Hu Li Zhong	교수	
9	黃賀寶 Huang He Bao	교수	hpwong@mx.nthu.edu.tw
10	王盈婷 Wang Ying Ting	교수	yingting.wang@utexas.edu
11	王鴻龍 Wang Hong Long	교수	hlw@mail.ntpu.edu.tw

자료: PSFD 홈페이지(조직도) https://psfd.sinica.edu.tw/web/plan_03en.htm에서 2018. 7. 31. 인출

□ 자문위원

〈표 6-9〉 PSFD 연구 자문위원

구분	이름	소개
국내	莊英章 Zhuang Ying Zhang	국립교통대학객가문화학원 교수
	瞿海源 Qu Hai Yuan	중앙연구원 사회소연구위원
국외	蔡瑞胸 Cai Rui Xiong	미국 시카고대학 통계학 교수
	Gary Becker	시카고대학 경제학 교수, 노벨상 수상자
	Angus Deaton	프린스턴대학 경제학 교수, NAS 회원
	Robert Hauser	위스콘신대학 사회학 교수, NAS 회원
	Ronald Lee	캘리포니아대학 인구학 교수, NAS 회원
	William Parish	시카고대학 사회학 교수
	George C. Tiao	시카고대학 통계학 교수, NAS 회원
	Cheng Hsiao	남캘리포니아대학 계량학 교수, NAS 회원

자료: PSFD 홈페이지(조직도) https://psfd.sinica.edu.tw/web/plan_03en.htm에서 2018. 7. 31. 인출

제4절 조사표의 주요 내용³⁰⁾

□ 대만 제1차 조사표

○ 교육 경험 수집(RCI2009부터)

- 초등교육부터 연구소까지 각 단계의 교육 경험 및 과정
- 과외학습 유무/학교정보/전공정보/유학 유무/학비 및 생활비의 부담 형태

○ 직업 경험

- 16세 때 부모의 직업과 거주 형태
- 첫 구직 시 직장 및 신분
- 졸업 후 첫 직장(2014년 이후)
- 현재 직업(겸직 포함): 최종 조사 내용과 동일, 매년 병역사항 조사

○ 결혼과 배우자

- 결혼 여부
- 결혼 전후 직업의 변화
- 결혼 당시의 상황: 원가족 독립 유무, 거주, 배우자를 만나게 된 경위, 원가족의 찬반 여부
- 배우자 기본 정보: 연령, 교육 정도
- 배우자의 현재 직업상황(겸직 유무): 매년 조사

30) 조사연구 데이터 아카이브(Survey Research Data Archive: SRDA) 홈페이지 https://srda.sinica.edu.tw/browsingbydatatype_result.php?category=surveymethod&type=2&csid=5에서 조사 종류별 조사표를 다운로드 가능. 2018. 7. 31. 기준으로 인출한 조사표를 최대한 자세하게 기술하였음.

- 친가(親家) 자료(부모, 배우자의 부모, 매년 조사)
 - 기본자료: 출생 연도, 교육 정도
 - 가장 오래 다닌 직장
 - 건강 상태
 - 동거 여부
- 형제자매
 - 기본자료: 성별, 나이, 교육 정도
 - 거주 형태
 - 직업 형태
- 거주 형태(매년 조사)
 - 거주지와 주변 현황
 - 동거 친지 여부
- 가족 유대관계 및 지출(매년 조사)
 - 집안일 소요시간
 - 정부 보조금 수취 현황
 - 각종 지출 현황
- 자녀교육
 - 자녀 수
 - 자녀 정보: 성별, 나이, 교육 및 돌봄 상황
 - 교육 및 양육 정도
 - 자녀정보는 매년 조사, 단 질문 내용에는 차이가 있음.

□ 대만 추적 주 조사표

- 조사표 코드: RII, RIII, ... (2004년 이전), RR(2005년 이후)
- 조사 대상: 주 표본, RCI 조사표의 주 표본 자녀
- 조사표의 주요 목적: 추적 방문 조사
- 주요 조사 내용
 - 생활 및 직업 형태
 - 혼인 및 배우자 자료(배우자의 직업정보)
 - 거주 형태(자녀정보, 부모, 배우자의 부모)
 - 가정 유대관계 및 지출 형태
 - 자녀의 교육 및 양육 형태
 - 직업 형태: 질문 넘어가기 설계[2012년 (포함)이전]
 - 목적: 조사 대상의 중복 조사 방지
 - 설계 논리: 조사 대상자의 현재 직업과 전 1차 방문 시 직업이 같을 경우 바로 다음 단계로 넘어가도록 설계

예) 현재의 직업이 지난 조사 시 직업과 같나요?

- ① 변했다.(A7a로 넘어가기)
- ② 이제 구직을 했다.(A8a로 넘어가기)
- ③ 직업이 바뀌었다.(A9a로 넘어가기)

- 고용주의 변화
- 직업 혹은 직위 변화
- 필수 질문: 회사 위치, 통근 소요시간, 월급, 근무시간
 - 2014년부터 직업에 대한 조사 내용에 변화가 있음.

- 직업이 있는 조사 대상자는 직업정보를 모두 기록함.
- 현재 직업의 시작일과 초기 월급, 근무시간 기록.
- 이직자: 이직 원인/이직 시기/이직 시 월급과 근무시간
- 혼인 및 배우자 자료 - 질문 넘어가기 설계

예) 지난번 방문일(**년 **월) 이후 혼인 형태에 변화가 있습니까?

(혼인/이혼/재혼)

① 변화가 없다. 혼인 상태 유지(A17로 넘어가기)

② 변화가 없다. 동거 유지(A29로 넘어가기)

□ 대만 자녀 조사표

- 조사표 코드: CI, CII, CIII 등
- 조사 대상: 주 표본 1965년 이후 출생자의 16~25세 자녀
- 조사의 주요 목적: 학력 및 부모와의 관계 수집
- 주요 조사 내용
 - 개인자료, 교육 경험
 - 직업 경험
 - 혼인 여부 및 배우자 자료
 - 거주 형태
 - 가정 결속력 및 지출

□ 동남중국 조사

○ 2004년 1차 조사

- 조사표: 대만지구 RI2003 조사표를 중심으로 수정 보완
- 주요 조사 내용
 - 개인자료
 - 교육 경험
 - 직업 여부와 연혁
 - 혼인 및 배우자 자료, 가정의 형태
 - 친인척 자료(부모, 배우자의 부모, 형제자매)
 - 주거 형태
 - 가족의 지출 자녀교육 및 양육 형태
 - 가족관계 및 친밀도
 - 가정 경제관리

○ 2006년 이후 조사

- 조사표: 대만지구 추적조사표를 중심으로 수정 보완
- 주요 조사 내용
 - 생활 및 직업 형태
 - 혼인 및 배우자 자료
 - 거주 형태(자녀와 부모 배우자 부모 현황 포함)
 - 가정 지출 사회관계 및 능동성
 - 가족관계 및 자녀 양육 형태
 - 가정 경제관리

□ 동남중국 조사표 세부 사항

○ 소득 및 지출

- 소득: 직업 항목에서 월 소득과 주거 형태에 따른 자산 여부
- 지출
 - 주택 대출금 월평균 비용
 - 보모 및도우미 월평균 비용
 - 식비 지출 월평균 비용
 - 기타(교육비, 지난해 총액/의료비, 지난해 총액/여행, 지난해 총액)
 - 매년 생명보험 및 의료보험 평균 총액
 - 부모가 토지 및 부동산을 증여했는지 여부

○ 복지 및 노동

- 직업
 - 현재 직업 유무
 - 회사명/직위/구체적 사항(기업 형태/임직원 수/위치)
 - 일일 근무시간/월급/근무 기간
 - 직업이 없다면 - 이유
 - 지난해 직업 여부
 - 지난해 근무지의 일일 근무시간/월급/근무 기간
 - 첫 직장을 구했을 당시 연도
 - 일일 근무시간/월급/근무 기간

□ 복지패널 조사표와 PSFD 조사표의 구체적 비교

○ 일반

- PSFD: 기본 개인 신상 외 흡연 및 음주 여부/주행 시 안전벨트 착용 여부 등
- 한국복지패널: 기본 개인 신상 외에 장애 유무, 장애등급, 맞춤형 급여 수급 형태 등

○ 소득 및 지출

- PSFD: 피조사자와 그 배우자의 월급을 소득으로 보고, 이 외의 주거 형태에 따른 자산을 산출. 월평균 생활비가 조사표에 포함됨(지출 세부 사항에는 생활비, 교육비, 사보험비, 주거비 등이 있음).
- 한국복지패널: 가구 구성원 모두의 소득을 직업별 급여액을 조사. 생활비 내역을 세분화하여 월평균 비용을 조사하고, 그 외의 사회보험 및 개인연금 가입 여부도 조사.

○ 복지 및 노동

- PSFD: 복지 관련 세부 조사가 없음. 직업에 대한 조사로는 구체적 직책 및 근무시간과 이직 여부 등
- 한국복지패널: 기초생활보장제도에 관한 세부 사항과 복지서비스 이용 여부가 구체적으로 포함됨. 직업에 대한 조사는 'PSFD'와 유사함.

○ 기타

- PSFD: 조사의 목적을 '중국인 가정의 경제, 사회, 심리, 관습 등에 대한 전반적인 연구'에 두고 있으며 조사의 세부 사항이

‘일반 중국 가정의 이해’에 중점을 두고 있음.

- 한국복지패널: 조사의 목적을 ‘한국 저소득층의 규모와 생활 실태 파악’에 두고 있으며, ‘소득, 복지, 보건’에 중점을 두고 있음.

제5절 패널 활용

□ 홈페이지 운영: <https://psfd.sinica.edu.tw/web/> 참조

○ 공고란: 주요 공고 내역

○ 기본 소개

- 조사 배경
- 조사 진행 형태와 미래전망
- 구성원 소개

○ 조사 활용 실적

○ 조사 구조 및 자료 소개

○ 자료 다운로드

- 자료를 다운로드하기 위해서는 조사연구 데이터 아카이브 (Survey Research Data Archive: SRDA) 홈페이지 (<https://srda.sinica.edu.tw/>)에서 회원 가입을 해야 함

○ 관련 기관

○ 연구원 연락처 및 웹사이트

□ 외부 자료 제공 방식

- 우선 SRDA에 회원 가입을 해야 함.
- 가입 방법
 - 홈페이지에서 회원 가입 신청을 하고, ‘회원약정조건동의서’를 인쇄하여 작성한 후 우편으로 발송한다(재직·재학 증명서 포함).
- 회원 유효기간
 - 일반회원: 2년
 - 임시회원: 1년
- 회원 의무사항
 - ‘회원 약정 조건’을 준수한다.
 - 회원은 자료를 다운받아 사용한 후 논저 등을 발표하고, 그 목록 및 연구 내용을 제공하여 타인의 연구에 도움을 주어야 함.

□ 활용 실적

〈표 6-10〉 PSFD 공식회원 - 저널 & 콘퍼런스 논문

Author	Year	Title	Publication
Hung Lin Tao Ching Chen Yin	2016	HEIGHT, PARTNERS AND OFFSPRING: EVIDENCE FROM TAIWAN	Journal of Biosocial Science, 48(5):593-615
Hung-Lin Tao	2014	Why do women interact with their parents more often than men? The demonstration effect vs. the biological effect	The Social Science Journal, 51(3), 350-360
Hung-Lin Tao	2014	Height, weight, and entry earnings of female graduates in Taiwan	Economics & Human Biology, 13, 85-98
C. Y. Cyrus Chu Seik Kim Wen-Jen Tsay	2014	Coresidence With Husband's Parents, Labor Supply, and Duration to First Birth	Demography, 51(1), 185-204
Hung-Lin Tao	2012	Informational Ambiguity and Survey Bias: Husbands' and Wives' Reports on Their Contribution to Their Families	Social Indicators Research, 111(3), 713-724
Ruoh-rong Yu Lung-An Li	2011	Imputation of Non-ignorable Nonresponses for Income: Analysis of a Panel Study on Taiwan	Quality & Quantity, 45(4), 875-884
C. Y. Cyrus Chu, Yu Xie, Ruoh-rong Yu	2011	Coresidence with Elderly Parents: A Comparative Study of Southeast China and Taiwan	Journal of Marriage and the Family, 73(1), 120-135
Tsai, Shu-Ling	2010	Language Skills and Status Attainment in Taiwan	Journal of Language, Identity, and Education, 9(4), 229-249.
Tsay, W-J	2009	The Fertility of Second-Generation Political Immigrants in Taiwan.	Journal of Immigrant & Refugee Studies 7, 109-128
C. Y. Cyrus Chu Ruey Tsay	2008	Intergenerational Transmission of Sex-specific Differential Treatments: The Allocation of Education	Social Science Research 37(2), 386-399

Author	Year	Title	Publication
R. Yu		Resources among Siblings	
C. Y. Cyrus Chu Yu Xie R. Yu	2007	Effect of Sinship Structure Revisited: Evidence from Intra-family Resource Transfer in Taiwan	Sociology of Education 80(2), 91-113
Wen-Jen Tsay	2006	The Educational Attainment of Second-Generation Mainland Chinese Immigrants in Taiwan	Journal of Population Economics, 19(4), 749-767.
Kamhon Kan	2006	Cigarette Smoking and Self-Control	Journal of Health Economics, 26(1), 61-81.
Kamhon Kan Wei-Der Tsai	2005	Parenting Practices and Children's Education Outcomes	Economics of Education Review, 24(1), 29-43.
Wen-Jen Tsay C.Y. Cyrus Chu	2005	The Pattern of Birth Spacing During Taiwan's Demographic Transition	Journal of Population Economics, 18(2), 323-336.
Kamhon Kan	2005	Parenting Practices and Children's Education Outcomes	Economics of Education Review, 24(1), 29-43
Kamhon Kan	2004	Obesity and Risk Knowledge	Journal of Health Economics, 23(5),907-934
Ruoh Rong Yu Yu Sheng Liu	2011	Changes in Gender Role Attitudes, Domestic Labor and Marital Power: Evidence from Repeated Cross-Sectional and Panel Surveys of Taiwan	Workshop on "Marriage in Cosmopolitan China", 香港
Ruoh Rong Yu Yu Sheng Liu Meng Li Yang	2011	Does Interviewer Personality Matter for Survey Outcomes? Evidence from a Face-to-Face Panel Study of Taiwan	World Association for Public Opinion Research (WAPOR), Amsterdam, Netherlands
Chang, Ying-Hwa Ruoh-Rong Yu	2008	The Persistence and Change of Family Structure in China and Taiwan: A Comparative Study of Residential Arrangement of the Aged	ISA Research Committee on Family Research (CFR), RC06, "Family Diversity and Gender," Lisbon, Portugal, September 9-13.
Chen, Chih-Jou Ying-Hwa Chang	2006	Self-Employment in Eastern China and Taiwan	Paper presented for Conference on Rethinking the Rural-Urban Cleavages in Contemporary China, Fairbank Center, Harvard University, Oct. 6-8, 2006.

자료: PSFD 홈페이지[(PSFD 데이터를 활용한 간행물) https://psfd.sinica.edu.tw/web/plan_04.htm#5에서 2018. 7. 31. 인출]

제6절 복지패널에 주는 시사점

□ PSFD와 한국복지패널의 공통점은 다음과 같음.

- 첫째, 대만에서 시작하여 중국의 3개 성(省)으로 확대해 진행되는 가구 단위의 패널추적조사로서 한국복지패널과 같이 규모가 큼.
- 둘째, 조사 내용을 바탕으로 다양한 분야의 학제 연구가 이루어지고 있으며, 무궁한 연구의 바탕이 될 수 있음.

□ PSFD와 한국복지패널의 차이점은 조사 목적에 있음.

- PSFD: ‘중국인 가정의 경제, 사회, 심리, 관습 등에 대한 전반적인 연구’에 목적을 두고 있어, 조사의 세부 사항이 ‘일반 중국 가정의 이해’에 중점을 둠.
 - ‘효도와 가정에 대한 관념’, ‘가족 간의 친밀도’, ‘자녀 교육 태도’ 등
- 한국복지패널: ‘한국 저소득층의 규모와 생활실태 파악’에 목적을 두고 있어, 조사의 세부 사항이 ‘소득, 복지, 보건’에 중점을 두고 있음.

□ ‘한국복지패널’에 ‘한국 저소득층 가정의 이해’와 관련된 세부 내역을 조사 내용으로 포함한다면 현재 이루어지고 있는 경제적 정책 형성과 제도적 개선 외에도 저소득층의 가족관계 개선을 위한 구체적인 사회복지제도 마련에 기여할 수 있을 것으로 기대됨.

제 7 장

결론

제1절 결론

제2절 한국복지패널에 주는 시사점

제1절 결론

- 한국복지패널 심층분석 보고서의 목적은 크게 두 가지로 볼 수 있음.
 - 첫째, 복지패널의 해외 홍보를 위해 외국의 주요 패널이 어떻게 구성되어 있으며, 자체 패널 홍보를 위해 어떠한 노력을 하고 있는지 파악하는 것임.
 - 둘째, 복지패널의 해외 홍보를 위해 우리가 가지고 있는 조사표 및 유저가이드를 영문화하는 데 목적이 있음.
 - 이 외에 연구 진행 과정에서 논의된 복지패널의 지속가능성 확보 방안에 대한 검토임.
 - 현재 1세대 분가만을 허용하고 있는 복지패널을 중장기적으로 2세대 혹은 3세대까지 분가를 확대한다면 어떻게 진행하는 것이 좋을지를 알아보기 위해 분가가구에 대한 추적규칙과 가중치 부여 방안 등을 분석하는 것임.
- 해외 주요 패널로는 조사 기간이 오래되고 활용도가 높은 미국의 PSID와 유럽 패널인 영국의 BHPS, 독일의 SOEP를 살펴보았으며, 아시아 국가의 패널로는 일본의 JHPS/KHPS와 대만의 PSFD를 분석함.
 - 분석 방식은 패널의 조사 배경, 표본 및 가중치, 조사 방식 및 관리, 조사표 주요 내용과 패널 활용을 중심으로 해외 패널에 대한 조사를 진행한 것임.

- 각 패널에 대한 조사 결과를 통해 우리 복지패널에 주는 시사점 등을 확인해 보면 아래와 같았음.

□ 영국 BHPS

- BHPS를 통해 확인할 수 있는 사항은 패널의 국제화와 활용도 제고를 위해 많은 노력을 하고 있었다는 점임.
- 패널의 국제화를 위해 주요국이 모여 패널 자료와 연구를 공유하는 CNEF에서 해당 패널을 공유하고 있었으며, 이 외에 격년 단위로 국제 세미나 등을 개최하고 있었음.
 - CNEF에서는 표준화된 방식으로 정리된 파일을 국제 홈페이지를 통해 공유함으로써 해외 전문가 및 학자 등이 쉽게 활용할 수 있도록 지원하고 있었음.
 - 자체적으로는 격년 단위로 BHPS 활용을 위해 국제 세미나 등을 개최하며 BHPS를 활용한 논문 발표와 방법론 등을 강의함으로써 연구자의 욕구에 부합하게 프로그램을 운영함.
- 자국 및 해외 패널 활용자를 위해서는 별도의 사용자 커뮤니티를 운영하며 패널 발전을 위한 하나의 자산으로 활용하고 있었음.
 - BHPS 사용 경험이 있는 연구자 등을 주요한 자산으로 관리하면서 패널 개선을 위한 의견을 청취하고 조사표 개발 등에 적극적으로 활용하는 방안 등을 모색하고 있었음.

□ 미국 PSID

- 미국의 PSID는 1960년대에 시작된 오래된 패널임에도 끊임없이 패널 확산과 우수성 제고를 위한 노력을 하고 있었음.

- 먼저 패널조사 관리 측면에서 조사 대상자, 조사원 등의 패널조사로 인한 부담을 줄이기 위해 응답시간에 대해 조사하고, 어려운 질문과 쉬운 질문 등을 구분하여 어려운 문항에 대상자들이 조금 더 쉽게 응답할 수 있는 방법을 고민함.
- 표본과 관련해서는 시기별로 중요한 패널표본을 추가하거나(이민자 패널) 분가가구에 대한 관리 강화로 패널표본이 줄어들지 않고 오히려 늘어나는 양상을 보여 줌.
 - 분가가구의 경우 1960년대 패널이 시작된 이후 약 4세대까지 조사가 이루어져 생애주기는 물론 세대 간 연구가 가능하다는 특성을 지니고 있음.
- PSID 역시 영국의 BHPS와 같이 CNEF에 정형화된 패널조사자료를 제공하고 여러 국가의 연구자들이 PSID를 적극 활용하도록 하고 있음.

□ 독일 SOEP

- 독일의 SOEP는 영국의 BHPS와 함께 유럽의 대표적 패널로서 독일의 현대사를 반영하듯 표본을 다양하게 구성하여 운영하는 특성을 지님.
 - 예를 들어 독일의 통일을 추진하며 동독 주민과 이후의 인구구조 변화를 반영하여 이민 및 난민 등을 포함하는 등 독일의 인구특성을 반영하기 위한 변화를 지속적으로 추진함.
- 인구구조의 변화와 함께 최근 독일의 생활상을 반영하여 주기적으로 패널의 조사표를 수정 및 보완함으로써 시의적절한 연구가 가능하도록 하고 있음.

- SOEP 역시 패널의 국제화를 위해 유럽국가들이 많이 참여하고 있는 LIS, LWS 등에 자료를 제공하고 있으며 이 외에 CNEF에도 관련 자료 등을 제공하고 있음.

□ 일본 JHPS/KHPS

- 일본 JHPS/KHPS는 일본의 대표적 패널로서 국제교류도 활발하게 진행하고 있는 패널조사 자료임.
 - 2017년 한국복지패널과도 한일 패널 교류를 위한 세미나를 서울에서 개최한 바 있음.
- JHPS/KHPS는 국제연계를 위해 LIS, CNEF 등에 자료를 제공하고 있으며, 이 외에 국제기구 등에 비교 연구 제공은 물론 관련 자료 제공도 하고 있음.
 - 자국민을 위해서는 연간 패널데이터 관련 세미나 및 워크숍 등을 운영함으로써 패널의 활용도를 높이려는 노력을 하고 있음.

□ 대만 PSFD

- 대만의 PSFD는 초기에 대만에서 시작되었지만, 이후 중국 본토까지 확대되는 과정을 거쳐 중국의 대표적 패널로 변모해 가는 과정에 있음.
- PSFD 역시 다양한 조사 내용을 포함함으로써 학제 간 연구가 가능하도록 노력하고 있으며, 중국 사회의 가족 등과 관련된 주요 변화상을 연구할 수 있도록 자료가 구축되고 있음.
- 반면 다른 비교 대상 패널과는 다르게 해외 패널과의 교류는 충분히 이루어지지 않고 있는 것으로 판단됨.

제2절 한국복지패널에 주는 시사점

□ 해외 패널을 통해 얻을 수 있는 공통적인 시사점을 정리하면 아래와 같음.

○ 첫째, 외부기관 등과의 협업으로 조사를 위한 펀딩을 확보하고 있다는 점임.

- 우리의 경우 매년 정부예산을 활용하여 조사를 위한 예산을 확보하는 것과 다르게 해외 패널의 경우 외부에서 펀딩을 지속적으로 받고 관련 기관의 욕구에 부합하는 연구는 물론 자체 연구를 계속해서 수행함으로써 서로 윈윈하는 방법을 택하고 있었음.

- 외부 자원, 기금 등을 활용하여 패널을 안정적으로 운영하는 것은 바람직하지만 우리나라의 여건상 외부 재원을 활용하는 것은 현실적으로 쉽지 않은 문제임.

- 기초과학, 기초사회과학 등에 대한 국가적 투자와 더불어 기업, 재단 등의 관심이 늘어나야 하지만 우리나라의 연구환경은 여기까지는 미치지 못하고 있는 것으로 보임.

- 하지만 장기적으로 패널과 관련된 데이터 설명회, 학술대회 혹은 별도의 콘퍼런스 등에 관련 기업이나 재단 등이 관심을 가질 수 있는 주제를 발굴하고 이를 활용한 연구 결과가 발표된다면 연구재원을 확보할 수 있는 방안의 하나로 검토될 수 있음.

○ 둘째, 패널 이용자를 관리하고, 이용자를 대상으로 다양한 콘텐츠를 제공하고 더불어 패널을 활용할 수 있도록 교육프로그램을 주기적으로 제공하고 있다는 점임.

- 우리나라의 경우 대부분의 패널이 학술대회를 중점으로 하여 이용자 관리를 하고 있지만, 해외 패널의 경우 패널을 활용할 수 있도록 교육프로그램을 운영하고 관련 경비를 지원하고 있음.
- 또한 이용자들이 쉽게 접근하고 패널을 활용할 수 있도록 패널 이용자 모임을 관리하고 관련 의견 등을 청취하는 모니터링을 하고 있음.
- 한국복지패널의 경우 오랜 기간 운영되고 있음에도 별도의 조직이 아닌 연구조직에서 관리하고 있어 실질적으로 이용자의 편의를 위한 서비스를 제공하는 것에는 한계가 있었음.
- 점진적으로 이용자 편의와 패널 활용도를 높이기 위한 방안으로 데이터 설명회와 학술대회를 콘퍼런스 혹은 워크숍의 형태로 일정 기간 운영하면서 이용자 모니터링을 동시에 하는 방안을 검토할 수 있음.
- 또한 우수 연구자 및 보고서를 발굴하여 관련 연구자에게 패널 활용 연구를 위한 연구비를 제공하고 이를 함께 발간함으로써 패널 활용도를 높일 수 있을 것임.

○ 셋째, 해외 패널과의 지속적 교류임.

- 영국, 독일, 일본, 미국 등의 주요 패널은 CNEF 혹은 자체적인 프로그램으로 해외 패널과의 지속적 교류를 진행함으로써 패널을 홍보하는 과정을 운영하고 있음.
- 한국복지패널의 경우 CNEF 혹은 해외 패널과의 교류 등을 모색한 바 있으며, 패널을 국제화할 필요성을 느끼고 있지만 이로부터 얻을 수 있는 실익에 대한 고민과 패널조직의 한계(인력, 전문조직 등)로 패널의 국제화가 충분히 진행되고 있지 못함.

- 필요에 따라서는 단계적으로 복지패널을 해외에 알리는 작업을 수행할 필요가 있음.
- 외국의 주요 연구자 혹은 국내에서 유학 혹은 파견 간 연구자 그룹에 연구비를 지원함으로써 좋은 연구 결과를 생산해 내는 것이 우선임.
- 점진적으로 국내 패널과 연합하거나 해외 패널과 연계해 공동 콘퍼런스 등을 개최함으로써 한국복지패널을 점진적으로 국제화해 나가는 방안을 추진할 필요가 있음.

□ 복지패널 중장기 표본 확보 방안

- 패널은 특성상 조사가 진행됨에 따라 패널 소실 혹은 표본 감소라는 문제에 직면하게 됨.
 - 복지패널을 비롯해 국내의 주요 패널은 패널 소실 혹은 표본 감소와 관련하여 주로 표본을 추가하는 방식을 취하고 있음. 하지만 이는 기존 패널과의 정합성, 가중치 등에서 혼돈을 유발할 개연성이 높음.
- 미국 PSID의 경우 분가가구를 지속적으로 추적조사함으로써 오히려 표본이 줄어들지 않고 늘어나는 효과를 가져오는 방식을 취하고 있음.
- 복지패널을 비롯한 국내 대부분의 패널이 1세대 분가만을 허용하고 있다는 점에서 중장기적으로 1세대 분가를 넘어서는 2세대 및 3세대로 넘어가는 분가가구에 대한 추적조사가 필요함.
 - 분가가구 조사의 장점은 기존의 1세대 패널과의 연계성을 가지고 있어 세대 간 연구가 가능한 것은 물론 조사자와의 관계

성이 확보되어 있다는 점에서 표본 감소를 최소화할 수 있음.

- 또한 표본을 추가적으로 투입하는 것에 비해 패널의 연속성을 담보할 수 있는 특징을 지닐 수 있음.
- 반면 분가가구가 늘어남에 따라 추가적인 예산이 소요된다는 점과 1세대 가구들이 자녀 분가가구에 대한 추가 조사에 난색을 표하는 경우가 많다는 점은 단점으로 지적할 수 있음.

□ 한국복지패널 심층분석 보고서를 통해 연구된 주요 연구 결과가 향후 한국복지패널의 발전에 주요한 밑거름이 되기를 바라며 패널이 오랜 기간 유지됨으로써 주요한 연구자료와 국가통계로서의 역할을 하기를 바랍.

참고문헌 <<

- 강석훈. 1997. 「유럽의 패널조사 현황과 시사점 - 가구 패널조사와 국제 비교학적인 관점을 중심으로」. 성신여자대학교·한국노동연구원
- 게이오대학 패널데이터 연구센터 홈페이지 <https://www.pdrc.keio.ac.jp/>
- 백지선, 심규호. 2013. “가구패널조사에서 횡단면 가중값 작성방법”. 통계개발원 2012년 하반기 연구보고서 1장
- 백화중, 김안나. 2004. 「빈곤 및 공공부조 패널 데이터 구축 방안」. 한국보건사회연구원 기초보장·자활정책 평가센터.
- 손창균(2012). 패널 가중치 조정 방안에 관한 고찰. 한국여성정책연구원 여성가족패널 Technical Reports.
- 송헌재. 2010. 재정패널의 대체표본 특성 분석. 제2회 재정패널 학술대회 발표자료집.
- 이상호. 2006. “서구 주요 패널의 조사전략 변화와 시사점”. 노동리뷰. 통권 제16호. 63~78
- 이지은. 2016. “성년을 맞는 노동패널 - KLIPS 20주년을 맞이하여”. 노동연구원. working paper(ppt)
- 이택면, 김승연, 우원규, 이주영, 강석훈. 2009. 2009년 여성가족패널조사 사업보고서 및 제2차 기초분석 보고서. 한국여성정책연구원 연구보고서.
- 일본가계패널조사 뉴스 제14호(2017.12). https://www.pdrc.keio.ac.jp/uploads/JHPSKHPS_panelnews_No.14.pdf (2018. 5. 15 인출)
- 주재선, 김영택, 송치선, 손창균, 임찬수. 2012. 2012년 여성가족패널조사. 한국여성정책연구원 연구보고서.
- 최항석, 박규범. 2015. 「수급자 패널구축의 유용성과 활용에 관한 연구」. 사회보장연구원
- 한국노동연구원, 각 연도. KILPS 1차년도 - 3차년도 USER's GUIDE.
- 한국보건사회연구원, 2017. 한국복지패널 & 일본가구패널 국제학술 세미나 자료
- 한국보건사회연구원, 서울대학교사회복지연구소, 2018. 한국복지패널 12차년

도 조사자료 유저가이드.

- Alexander, C. H., S. Dahl, and L. Weidmann. (1997). "Making Estimates From the American Community Survey." Paper presented to the Annual Meeting of the American Statistical Association(ASA), Anaheim, CA, August 1997.
- Anne Bohlender, Simon Huber, Axel Glemser (Kantar Public). 2018. SOEP-Core - 2016: Methodenbericht Stichproben A-L1. SOEP Survey Papers 493: Series B. Berlin: DIW/SOEP
- American Community Survey Documentation Memorandum Series #ACS05-DOC-2. Washington, DC: U.S. Census Bureau.
- Bates, L. (2008). Editing the MAF Extracts and Creating the Unit Frame Universe for the American Community Survey. 2009 American Community Survey Universe Creation Memorandum Series #ACS09-UC-1. Washington, DC: U.S. Census Bureau.
- Berglund, P. A., Chang, W., Heeringa, S. G., & Mcgonagle, K. (2017). Panel Study of Income Dynamics PSID Cross-sectional Individual Weights, 1997-2015. Retrieved from https://psidonline.isr.umich.edu/data/weights/cross_sec_weights_15.pdf
- British Household Panel Survey(BHPS) Questionnaires (wave 18)
- Biemer, P., and L. Lyberg. (2003). Introduction to Survey Quality. Hoboken, NJ: John Wiley & Sons, Inc.
- DIW · Berlin. About SOEP. https://www.diw.de/en/diw_02.c.221178.en/about_soep.html. (2018. 5. 1 인출)
- Duffy, D. (2011). 2007 PSID Income and Wage Imputation Methodology. Survey Research Center - Institute for Social Research, University of Michigan. Retrieved from https://psidonline.isr.umich.edu/publications/Papers/tsp/2011-03_Duffy.pdf

- Duncan, G.(1995), "A Simple Method for Weighting in Household Panel Survey", Working paper, Northwestern University.
- Earnst, L.(1989), "Weighting Issues for Longitudinal and Family Estimates", In Kasprzyk, D. Duncan, G., and Singh, M. eds, Panel Survey, Wiley, pp.139-177.
- Fay, R. (2006). Using Administrative Records with Model-Assisted Estimation for the American Community Survey. JSM Proceedings, Survey Research Methods Section (pp. 2995-3001). Alexandria, VA: American Statistical Association.
- Gouskova, E., Heeringa, S. G., McGonagle, K., Schoeni, R. F., & Stafford, F. (2008). Panel Study of Income Dynamics revised longitudinal weights 1993-2005. Ann Arbor: University of Michigan. Retrieved from https://psidonline.isr.umich.edu/publications/Papers/tsp/2008-05_PSID_Revised_Longitudinal_Weights_1993-2005%20.pdf
- Groves, R. M., F. J. Fowler, M. P. Couper, J. M. Lepkowski, E. Singer, and R. Tourangeau, (2004). Survey Methodology. Hoboken, NJ: John Wiley & Sons, Inc.
- Institute for Social Research. (2017). PSID Main Interview User Manual: Release 2017. University of Michigan, May, 2017. Retrieved from <https://psidonline.isr.umich.edu/data/Documentation/UserGuide2015.pdf>
- Institute for Social Research (1972), A Panel Study of Income Dynamics, University of Michigan.
- Kalton, G., and Brick, M.(1994), "Weighting Schemes for Household Panel Surveys", Proceedings of the Section on Survey Research Methods, American Statistical Association, 789-790.
- Latouche, M., Dufor, J., and Merkouris, T.(2000), "Corss-sectional

- Weighting: Combining Two or More Panels”, Statistics of Canada Research Paper series.
- Lecvesque, I., and Franklin, S.(2000), “Longitudinal and Cross-Sectional Weighting of Survey of Labour and Income Dynamics 1997 Reference year”, Statistics of Canada Research Paper series.
- McGonagle, K. A., and Schoeni, R.F.(2006), “The Panel Study of Income Dynamics : Overview and Summary of Scientific Contributions After Nearly 40 years”, Institute for Social Research, University of Michigan.
- Mcgonagle, K., Schoeni, R., Sastry, N., & Freedman, V. (2012). The Panel Study of Income Dynamics: Overview, Recent Innovations, and Potential for Life Course Research. Longitudinal and Life Course Studies, 3(2), 268-284. Retrieved from <https://psidonline.isr.umich.edu/llcs2012.pdf>
- Naud, J. F.(2004), “Combined-panel longitudinal weighting Survey of Labour and Income Dynamics”, Statistics of Canada Research Paper series.
- PSID Homepage. <https://psidonline.isr.umich.edu/>
- SOEP. SOEP-Core. <https://data.soep.de/soep-core> (2018. 4. 17 인출)
- Survey Research Center, Institutue for Social Research, University of Michigan.. <http://www.src.isr.umich.edu/about/>
- Survey Research Operations. <http://www.src.isr.umich.edu/services/survey-research-operations/>
- Taylor, Marcia Freed (ed). with John Brice, Nick Buck and Elaine Prentice-Lane (2010) British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices. Colchester: University of Essex.
- TAXSIM program. <http://users.nber.org/~taxsim/to-taxsim/psid-kks/>

- Wager, Gert G., Frick, Joachim R., Schupp, Jürgen. 2007. "The German Socio-Economic Panel Study (SOEP) - Evolution, Scope and Enhancements" SOEPpapers on Multidisciplinary Panel Data Research. Schmollers Jahrbuch 127(1). (http://www.diw.de/documents/publikationen/73/diw_01.c.60184.de/diw_sp0001.pdf)
- Watson, N. (2004). "Wave 2 Weighting", HILDA Project Technical paper Series No. 4/04.
- Watson, N. and T. Fry(2002). "The Household, Income and Labor Dynamics in Australia (HILDA) Survey : Wave 1 Weighting", HILDA Project Technical paper Series No. 3/02.
- Weidman, L., Tsay, J., & Ikeda, M. (2007). Comparison of Alternatives for Controlling Group Quarters Person Estimates in the American Community Survey. Statistical Research Division Research Series RRS2007-4. Washington, DC: U.S. Census Bureau.
- Westat(2001), Survey of Income and Program Participation User's Guide, 3rd eds.
- 野村総合研究所, 2013. 日本におけるパネルデータの整備に関する調査 報告書
- 石井 加代子, 野崎 華世, 2014. 「慶應義塾家計パネル調査(KHPS)」と「日本家計パネル調査(JHPS)」におけるCross-sectional/Longitudinalウェイトおよびパネル統合ウェイトの作成. 三田商学研究 57(4), 123-145
- 黃奕嘉(황혁가), 2017. 華人家庭動態資料庫: 訪問架構與問卷設計(화인가정동능자료고: 방문가구여문권설계)
- 于若蓉·章英華(우약용·장영화), 2014. 華人家庭動態資料庫: 台灣訪問計畫(‘화인가정동능자료고: 대만방문계획)
- 家庭動態資料庫的建立 第十六年計畫(가정동태자료고적건립 제16년계획) CIX 2016 問卷(문권): PSFD 2016년 CIX 조사표

부록: 복지패널의 분가관리 << 방안을 위한 추적규칙과 가중값 조정에 대한 연구

제1절 패널 표본의 추적규칙

가구 패널조사는 개인 정보뿐만 아니라 가구 정보를 동시에 얻기 위한 목적으로 수행한다. 가구에 대한 정보를 수집하기 위해 가구 패널은 면접 가구에 속한 모든 개인을 포함하도록 설계되어야 하는데, 가구 구성원들은 시간이 지나면서 구조가 복잡해지기 때문에 이러한 사실을 밝히는 데 많은 노력이 요구된다. 이러한 가구 구성의 변화를 현장에서 어떻게 다룰 것인가에 대한 명확한 규칙, 즉 추적규칙(follow-up rule)이 필요하다.

패널 조사에서 추적규칙은 표본 가구 또는 개인이 패널에서 유지될지 또는 제외될지에 대한 상황을 결정하기 때문에 표본 변화의 중요한 원인이 되며 조사가 진행될수록 무응답, 표본 탈락 등의 도전에 직면한다. 즉, 추적규칙은 가중치 추정에 영향을 주는 중요한 요소로서 무응답과 표본 탈락의 영향을 줄이기 위해 가중치 조정은 필수적인 작업이며, 이를 위해 보정 과정(calibration approach)이 활용된다. 이 글에서는 해외 주요 패널조사에서 수행하는 추적규칙과 가중값 추정, 무응답 조정 방식, 보정의 영향 등에 대해 살펴보려고 한다.³¹⁾

전술한 바와 같이 가구 패널조사는 표본으로 추출된 개인이 사망 또는 이민 등으로 목표 모집단에서 탈락할 때까지 지속적으로 추적조사를 하게 된

31) 이 글은 복지패널의 분가가구를 추가적으로 조사하면서 발생하는 가중값 조정 문제에 대한 연구 결과이다. 초고에서는 가중값에 대한 이론적 고찰부터 해외 주요 패널에서 각각 가중값을 부여하는 구체적인 측정산식까지 제시하였으나, 독자의 이해를 돕기 위하여 전체 내용을 요약하여 수록하였다.

다. 표본으로 추출된 개인은 초기 추출확률을 보유하며, 이를 통해 안정된 추정치를 산출한다. 동거인에 대한 정보는 초기에 표본으로 포함된 개인에 대한 가구정보를 통해 수집하게 된다. 결과적으로 개인에 대한 모수를 추정할 때 이와 같은 동거가구원의 개인정보는 생략될 수 있다.

비싼 유지비용이 투입되는 패널조사에서 초기에 표본으로 추출된 개인 자료만 이용하는 것은 비효율적이다. 경우에 따라 일부 패널조사는 초기 웨이브 이후에 면접한 모든 개인과 접촉한다. 이러한 패널조사의 경우 초기에 표본으로 포함되지는 않았지만 표본으로 추출된 개인과 같이 살고 있다는 사실만으로 면접을 실시한다. 이러한 패널조사 방법을 사용하는 이유는 이 사람들과 접촉하기에 앞서 추적이 상대적으로 용이하며 따라서 현장에서 직접 표본으로 추가하는 것에 비해 상대적으로 적은 비용이 소요되기 때문이다. 초기 표본을 구성할 때 되도록 많은 개인을 추적해야 하는 또 다른 이유는 표본 탈락과 사망으로 발생하는 부정적 효과를 제거하기 위함이다. 또한 표본에 포함된 부모 패널에게서 태어난 신생아에 대한 규칙을 만들어야 한다. 이 경우 초기에 아동이나 성인에 관한 정보를 수집해야 한다. 현재 수행되는 다수의 패널조사에서는 이러한 개인들을 표본에 포함하도록 하는 일련의 추적규칙을 가지고 있다.

패널조사의 목적은 시간의 흐름에 따라 이러한 개인을 추적하여 조사하는 것이다. 그러나 종단면적 설계에서 특정 시점 t 에서의 연구 모집단을 정의하는 것은 쉬운 일이 아니다. 이러한 모집단은 초기 모집단, 동거자를 포함한 초기 모집단, 자식세대를 포함한 초기 모집단, 가구관계에 기초한 친족 모형, t 시점까지 횡단 모집단의 일부로 모든 개인을 고려한 모집단 등으로 가정할 수 있다.

먼저 초기 모집단은 시간의 흐름에 따라 모집단이 생성되며, 모집단 프레임에서 사망 및 이주는 추정에서 제거된다. 이러한 결과 초기 모집단은

시점 t 에서의 횡단면적 모집단과 차이가 발생하게 된다. 동거자를 포함한 초기 모집단은 t 시점에서의 모집단에 초기 모집단 가구에 동거하고 있는 모든 사람을 포함한다. 자식세대를 포함한 초기 모집단은 t 시점의 모집단에 있는 자손들을 포함하는 초기 모집단의 개념이다. 결과적으로 이러한 모집단은 초기 모집단이 인구학적으로 확대된 형태이다. 이러한 개념은 미국의 PSID와 영국의 BHPS, 유럽연합의 ECHP에서 도입하였다.

가구관계에 기초한 친족 모형은 초기 모집단에 속한 개인과 함께 살고 있는 친족이 있다면 이들을 초기 모집단에 포함하는 것이다. 이것은 보다 높은 차원의 관계로 정의할 수 있으며, 독일의 SOEP가 이러한 방법을 사용하고 있다. t 시점까지 횡단적 모집단의 일부로 모든 개인을 고려한 모집단은 시점 0에서 시점 t 까지의 모든 횡단적 모집단을 하나로 통합하는 모집단으로 고려하는 것이다. 이러한 모집단에 대한 통계를 산출할 때는 적절한 결측치 처리가 필요하다.

표본은 시간의 흐름에 따라 새로운 개인을 포함하게 되며, 이러한 표본은 초기 표본과 차이가 발생한다. 조사가 지속될수록 이러한 개인을 어떻게 추적하고 표본에 포함할지에 대한 추적규칙이 필요하다. 이러한 추적규칙에는 초기 표본에 대해서만 추적하는 규칙, 초기 표본과 동거자만을 추적하는 규칙, 초기모집단과 자손을 추적하는 규칙, n 차원 동거인까지 추적하는 규칙 등이 활용되며, 각각의 규칙은 다양한 패널조사에서 조사 환경에 따라 다르게 적용되고 있다.

제2절 패널의 가중값 부여 방법

1. 가중값 부여 방안에 대한 이론적 고찰

가중값 공유 방안은 가중값을 계산하기 어려운 비표본가구원의 가중값을 원시표본가구원의 가중값과 공유하는 방안이다. 이러한 개념은 Ernest(1989), Lavalley(1995), Kalton and Brick(1995), Rendtel and Harms(2006) 등에 의해 점차 발전되어 왔다.

먼저 가중값을 공유하는 통상적인 형태는 원표본가구원에게만 가중값을 부여하고 비표본가구원에게는 0의 가중값을 부여하는 방법이며, 원표본가구원만을 추적하는 패널조사에서 사용이 가능하다. 그러나 이러한 패널 추적규칙에서는 표본 개인의 사망으로 표본의 규모가 감소하는 패널이 되며, 만약 원표본가구원과 동거하는 비표본가구원도 조사하였다면 이들의 정보를 상실하게 된다.

다음으로 표본 가구에 속하였던 모든 개인(표본가구원과 비표본가구원)을 추적조사하는 경우, t 시점의 가구 구성원에게 가중값을 모두 균등하게 배분하는 방법을 사용할 수 있으며 이것은 BHPS에서 활용하는 방법이다.

한편 또 다른 방법으로 가구 내 동일 가중값 공유(Equal Household Share) 방안에 대해서도 생각해 볼 수 있다. 이것은 동등개인가중값은 원시표본가구원과 비표본가구원의 개인추출확률이 동일하다고 가정하는 것과 같으며, 동등가구가중값은 원시표본가구와 비원시표본가구의 가구추출확률이 동일하다고 가정하는 것과 의미가 같다.

또 다른 방법으로 룩셈부르크의 가구패널조사인 PSELL(Panel Socio-Economic in Luxembourg)에서 활용하는 방안인 성인 가중값 공유(weight share of adults) 방안이 있을 수 있다. 이는 패널조사에서

의 추적규칙, 즉 종단면 모집단이 $t-1$ 시점의 성인으로 한정된 경우에 가중값 공유 방안을 활용하는 것이다. 이 외에도 PSID에서 사용되는 방식으로, 성인과 그들 사이에서 태어난 자녀의 경우 성인 또는 자녀들의 가중값 공유 방안을 사용할 수 있다. 이는 SOEP에서 활용한 방식이다.

2. 외국 패널의 가중값 부여 방안³²⁾

가. BHPS의 가중값 부여 방안

BHPS에는 다양한 형태의 가중값이 제공되고 있으며, 가중값 계산 과정에서 세세한 부분까지 조정하는 특징을 가진다. 가중값은 일반적으로 응답개인가중값, 기록개인가중값, 가구가중값으로 구성되어 있으며, 이는 다시 종단면 분석(longitudinal analysis)과 횡단면 분석(cross-sectional analysis) 등 분석 목적에 따라 구분된다.

종단면 분석에 사용되는 가구가중값은 2차 웨이브부터는 산출하지 않고 있는데, 이는 종단면 분석의 관점에서 시간에 따라 가구 구성원이 지속적으로 변화하므로 같은 가구로 보기 어렵기 때문이다.

원표본 응답가구원에 대한 개인 종단면 분석은 최근 웨이브의 종단면 응답개인가중값을 이용하여 분석한다. 아동을 포함하는 모든 원표본가구원의 개인 종단면 분석은 최근 웨이브의 종단면 기록 개인가중값을 사용하여 분석한다. 모든 응답개인(비표본가구원 포함)의 횡단면 분석에는 해당 웨이브의 횡단면 응답개인가중값을 활용한다. 아동과 비표본가구원을

32) 외국 패널(영국의 BHPS, 호주의 HILDA, 미국의 PSID, 캐나다의 SLID)에서 각각 종단면 가중값과 횡단면 가중값을 어떻게 계산했는지에 대해서는 이택면 외(2009)과 손창균(2012), 주재선 외(2012)에서 자세히 밝힌 바 있다. 이하 글은 두 연구에서 주요 내용을 인용하였으며, 최근 변화 내용 등은 각 패널조사 연구진이 발간한 자료를 참고하였다.

포함하여 모든 개인의 횡단면 분석은 해당 웨이브의 횡단면 기록개인가중값을 사용하여 분석한다.

〈부표 1〉 BHPS에서 가중값의 종류

구분		가중값의 종류	변수명
횡단면 분석	1차 웨이브	응답가구원가중값	AXRWGHT
		기록가구원가중값	AXEWGHT
		가구가중값	AHHWGHT
	2차 웨이브 이후	응답가구원가중값	wXRWGHT
		기록가구원가중값	wXEWGHT
		가구가중값	wXHWGHT
종단면 분석		응답가구원가중값	wLRWGHT
		기록가구원가중값	wLEWGHT

자료: 이택면(2009). p.77, 재인용.

나. HILDA의 가중값 부여 방안

1) 1차 웨이브 가중값 작성 방안

HILDA의 전반적인 가중값 부여 과정은 다음과 같다. 먼저 가구가중값의 경우 표본추출을 반영한 설계가중값을 기초로 한다. 이 가중값에 조사를 진행하면서 수집된 가구의 정보를 이용하여 가구의 무응답을 반영한 중간 단계의 가중값을 산정한다. 이 가중값에 호주 전체 모집단가구에 대한 정보를 이용하여 최종가구가중값을 산정한다.

기록가구원가중값은 최종가구가중값을 기초로 작성한다. 이 가중값에 호주 전체 모집단 개인에 대한 정보를 이용하여 최종기록개인가중값을 산출한다. 응답개인가중값도 최종가구가중값을 기초로 한다. 이 가중값에

수집된 가구원 정보를 반영하여 무응답조정가중값을 산정한다. 이 가중값에 호주 전체의 모집단 정보를 반영하여 최종응답개인가중값을 산출한다.

2) 2차 웨이브에서의 가중값

2차 웨이브부터는 총 다섯 가지의 가중값이 작성되는데 ① 2차 웨이브 가구분석을 위한 횡단면 가구가중값, ② 2차 웨이브 기록개인의 횡단면 분석을 위한 횡단면 기록개인가중값, ③ 2차 웨이브 응답개인의 횡단면 분석을 위한 횡단면 응답개인가중값, ④ 1, 2차 웨이브에 모두 존재하는 기록개인을 위한 종단면 기록개인가중값, ⑤ 1, 2차 웨이브에 모두 응답한 응답개인을 위한 종단면 응답개인가중값이 그것이다.

먼저 2차 연도 횡단면 개인가중값 작성 시 가구가중값을 먼저 산출하고 그 다음 개인 단위의 횡단면 가중값을 구한다. 2차 연도 횡단면 가중값 사후층화는 2차 연도 모집단 정보를 이용한다. 기록개인과 응답개인의 사후층화변수는 동일한 변수를 사용하며, 다만 기록개인은 모집단 개인 전체에 사후층화를 적용하고, 응답개인은 모집단 중 16세 이상의 경우만 적용하여 사후층화한다. 응답개인의 종단면 가중값을 작성할 때는 가구 단위의 무응답을 조정하고 이를 바탕으로 중간 단계의 가중값을 작성한 후, 개인 단위의 사후층화를 실시한다. 기록개인의 종단면 가중값은 개인 단위의 무응답을 조정한 후 개인 단위의 사후층화(16세 이상의 경우)를 적용한다. 기록개인과 응답개인의 경우 모두 사후층화는 1차 연도 모집단의 정보와 일치시킨다.

HILDA는 처음 조사 대상으로 선택된 가구원, 해당 가구에서 태어나거나 입양한 아동, 가구 구성의 변화로 새롭게 추가된 가구 구성원 등을 조사 표본으로 추적한다. 그 밖의 모든 신규 진입자는 임시거주자로서 추적

조사에서 제외되지만 만약 신규 진입자가 최초 표본 사이에서 아이를 가지게 되면 이 신규 진입자는 비표본가구원에서 표본가구원으로 변화되며, 이들은 지속적인 추적 대상이 된다. 다른 추적원칙은 일반적인 PSID 추적원칙을 따른다.

다. PSID의 가중값 부여 방안

1) PSID의 추적원칙과 표본 개인 접촉

1차 웨이브 면접 이후 가구(family) 단위에서 이탈한 모든 개인을 포함하여 PSID 표본 단위들을 추적하도록 한다. 가구원에 대한 PSID의 추적원칙은 가중값 산정과 한 시점의 전국 국민을 대표할 수 있도록 설계되었다. 원표본에서 태어난 신생아와 원표본의 입양아는 표본 단위로 분류하였고, 이들의 가구원으로 들어왔을 때 분가한 가구들을 추적하기 위한 적격 표본으로 고려하였다. 분가가구로 분류하기 위해 재면접가구로 분류된 기타 가구원을 포함하여 이들 개인은 ‘분가(split-off)’로 고려하고, 이들의 신규 가구원을 추적하였다. 이러한 과정은 인구집단의 가구 구성활동이 반복적으로 일어난다는 것과 매년 가구 표본의 동적 변화를 나타내게 된다.

아동이 성인으로 성장하여 신규 가구를 생성하거나 결혼으로 인해 새로운 PSID 가구가 지속적으로 생성된다. 분가가구의 추가적인 발생과 높은 조사완료율로 인해 연구 기간 동안 표본은 확대된다.

PSID의 추적원칙을 이해하기 위해서는 표본개인(sample person)의 개념을 이해해야 한다. 1993년까지 표본개인을 원표본 개인과 조사기간 동안 조사에 참여했던 표본 개인의 신생아와 입양아로 정의하였다. 어린

이는 출생 시점에 따라 연구기간에 일부 나타났다. 이러한 정의는 1994년까지 연구기간 동안 조사에 참여하지 않은 표본개인의 신생아와 입양아로 확대하였다. 이 경우 어린이는 출생시점에 응답 패널 가구로 이주할 필요가 없다.

주요한 PSID 추적규칙은 이전 웨이브에서 응답한 표본 개인을 적극적인 추적 대상으로 고려한 것이다. 1992년까지 18세 이하 개인의 표본 단위들은 이들의 가구로부터 이주하지 않았더라도 이들의 의견에 따라 추적하지 않았다. PSID는 1993년 이러한 규칙을 완화하여 18세 이하의 가구원들을 추적하였고, 이들이 속한 신규 가구에 대한 면접을 실시하였다.

PSID 표본 가구는 항상 가구원의 가구주(head) 또는 배우자(wife)를 포함한다. 그러나 이러한 경우 대상 연령 이하의 가구원이 비표본 부모로 편입되면 추적이 불가능하게 된다. 그러므로 1993년부터 대상 연령 이하의 가구원을 계속 추적하기 위해 가구주(head)와 배우자(wife) 둘 다 비표본 개인으로 조사된다.

이 경우 PSID는 특정한 연구 영역을 위해 비표본 응답자 그룹을 대상으로 면접을 실시하였다. 첫째, '비표본 노인'으로 65세 이상을 1990년과 1995년에 추적하였다. 둘째, 어린이 표본의 '비표본 부모'로 1994년에 추적을 시작하였으며, 웨이브마다 지속하다가 2005년에 종료하였다. 결과적으로 이러한 비표본 응답자는 1990~2003년 PSID 가구 표본에 표본 단위로 포함되지 않는다.

2) PSID 표본 추출과 가중값의 계산

가중값 계산을 위해 PSID의 표본추출 과정을 이해할 필요가 있다. 1968년 당시 가족표본은 다음 두 가지로 구성됐다. (1) 공통적으로 미국

에 거주하는 횡단면 표본과 (2) 통계국(Census Bureau)에 의해 OEO(Office of Economic Opportunity)에 대해 1967년에 조사된 가족들의 부차표본이다.

1969~1970년 표본은 전년도에 이어 계속 조사 대상이 되는 가구의 모든 가구원으로 구성된다. 따라서 전년도 웨이브에 응답하지 않은 구성원에 대해서는 2차 연도 조사를 하지 않았다. 거처의 횡단면 표본을 SRC(Survey Research Center)의 마스터 프레임으로부터 추출하였다. 마스터 표본은 설계에 유연하기 때문에 1개 이상의 거처를 표본으로 추출하거나 서로 다른 필요성을 가진 조사에 대해 적절한 시점에서 사용이 가능하다. 이 조사에서 3,000개의 거처가 조사될 수 있도록 설계되었다. 1968년의 센서스 표본은 재면접과 같은 형태인데, 이는 이들 가족들에 대해 통계국의 전년도 조사가 이루어졌기 때문이다. 여덟 가지의 기본적인 추출률을 가진 확률 표본추출이지만 통계국에 의해 조사된 가족들 중 소득이 2,000달러+N(1,000달러) 이하인 가족에 대해서만 조사되었다(여기에서 N은 가족 인원수를 의미). 2,000달러+N(1,000달러)값은 1967년에 사용된 연방 빈곤선(federal poverty line)의 2배와 거의 같다. 이 값 이상인 소득을 갖는 가족은 제외하였으며 북동부, 북중부, 서부의 3개 지역에 있는 표준 대도시 통계 지구(Standard Metropolitan Statistical Areas: SMSA) 외부 지역으로 빈곤가족에서 제외하였다.

각각의 표본은 1968년 조사에서는 무응답이었다. 이는 재면접 표본들이 인구센서스 조사에서부터 무응답자들이었기 때문이다. 즉 이들은 통계국에 의해 조사된 응답자의 이름과 주소를 OEO에 공개하는 것에 대한 서명을 거부하였으며, OEO로부터 SRC에 일부 표본 주소를 전달하는 데 실패하였다. 이에 따라 초기 가중값을 결정하기 위해 ① SRC 횡단면 표본에서 얻은 확률, ② 재면접 표본에서 얻은 확률, ③ 결합된 표본에서 얻은

확률을 각각 계산하였다.

PSID의 종단면 가중값³³⁾은 포함확률의 역수로 산정된 가중값이다. PSID 가구와 개인가중값에 대한 연도별 가중값 조정은 두 단계로 이루어진다. 먼저 개인 종단면 가중값을 조사 기간 내 조사에 참여한 모든 개인에게 부여하고, 두 번째로 조사에 참여한 모든 가구원의 개인가중값의 평균으로 가구가중값을 부여한다. 이때 표본 개인과 비표본 개인 모두에 대한 가중값의 평균을 계산한다.

개인가중값의 산정은 두 단계로 이루어진다. 첫째, 표본은 2개의 연속된 웨이브에서 개인의 상태에 따라 층을 나눈다. 가중값 산정 과정에서 항상 2개 웨이브를 고려하는데, 이는 ‘이전 연도’와 ‘현재 연도’를 참고하기 위해서이다. 다음으로 두 번째 단계에서는 각각의 개인에 대해 개인이 속한 층에 따라 개인가중값을 부여한다.

라. 미국의 SIPP(Survey of Income and Program Participation)

1970년대 후반에 미국 보건교육복지부(Department of Health, Education and Welfare: HEW)가 처음으로 소득조사(Income Survey Development Program: ISDP)를 시작하였다. ISDP의 내용과 과정의 발전으로 HEW는 조사표의 길이, 조사 기간의 길이, 조사 자료의 연계 프로그램 등에 관심을 가지게 되었다. 1979년 ISDP 패널은 종단면 조사였으며, 응답자는 그들의 소득과 노동시장 참여, 다른 특징에 관해 응답하도록 하였다. 이때 응답자들은 그들과 그 외 친인척의 정보를 제공하기 위해 3개월마다 접촉했다. 여기에서 3개월 간격은 면접에 대한 조사 주기

33) Gouskova, E. et al(2008). 중 ‘Methodological Approach to the Longitudinal Weight Construction’ 내용을 참고함.

(reference period)이다.

SIPP의 주된 목적은 미국의 개인 및 가정의 소득 현황, 분배 정책에 관한 정확하고 포괄적인 정보를 제공하는 데 있다. SIPP는 세금, 자산, 부채, 정부 지원 프로그램 참여 등에 관한 데이터를 수집한다. SIPP 데이터는 연방정부, 주정부, 지방정부 프로그램의 효율성을 평가하는 기초 자료가 된다. SIPP는 1996년 재설계되어 중요한 변화를 주었다. 우선 1996년 패널은 4년의 간격과 12개의 웨이브로 구성된다. 이러한 재설계는 초기 SIPP의 중복(overlapping) 패널 구조를 버리고 표본 규모는 증가시켰다. 1996년 초기 표본 규모는 4만 188개 가구였다. 각각의 SIPP패널에 있는 성인들은 민간인인 미국 모집단에서 가구들의 대표 표본에서 뽑힌 것이다. SIPP 표본으로 뽑힌 사람들은 패널 생애 동안 4개월에 한 번씩 조사에 응하게 된다. 만일 15세 이상인 원표본 구성원이 원래의 주소에서 다른 주소로 이사를 갔다면 새로운 주소에서 그들을 조사한다. 조사 대상 표본은 원표본 구성원과 함께 거주하는 어린이들을 포함한다. 만일 초기 조사 후 이전 조사에 없었던 사람들이 응답자 가구의 일부가 되었다면 새로운 사람들도 최초 조사한 응답자와 계속 살고 있는 동안 조사한다.

SIPP는 복합 표본 설계를 사용하고 있으며, 통계국에서 2단계 표본 설계에 의해 SIPP 표본을 추출한다. 1단계에서는 PSU를 선택하고 2단계에서는 표본 PSU에서 주소를 추출한다. 통계국 면접자는 선택된 주소에서 표본 구성원으로 식별하는 과정을 밟는다.

표본 PSU의 추출을 위한 프레임은 미국의 군(county) 지역과 독립 시(city)의 리스트로 구성된다. 이와 함께 모집단 수(counts)와 가장 최근의 모집단에 대한 센서스 결과에서 얻은 기타 자료를 함께 사용한다. 군(county)은 PSU의 형태를 갖도록 인근 군들로 그룹을 만들어 PSU로 하거나 하나의 군 자체가 PSU가 된다.

PSU의 구성 과정은 적은 PSU의 경우 인근의 비슷한 군들로 그룹을 이루어 하나의 PSU를 이루게 되는데 이를 non-self-representing(NSR) PSU라 하며 층을 구성한다(남부, 북동부, 중서부, 서부). 이 과정에서 인구 학적인 변수나 사회경제학적인 변수들이 최적의 그룹화를 위해 사용된다. NSR PSU들 중 하나의 표본을 층에서 모든 PSU를 대표하도록 각 층에서 추출한다. 특정 기준 이상의 큰 규모의 모든 PSU는 표본에 포함되며 이를 self-representing(SR) PSU라 한다. 통계국에 의해 유지되는 5개의 분리된 중복되지 않는 추출 틀에서 주소를 추출한다. 이들은 하나의 단위를 구성하며 주소 조사구 틀(address enumeration districts)이라 한다. 즉 ① 지역 프레임, ② 그룹쿼터 프레임(special places 프레임), ③ 가구 단위 포함 프레임, ④ 포함률 개선 프레임, ⑤ 신축건물(new-construction) 프레임 등이다. 이때 ①, ②, ③번 프레임은 센서스에 기초한 프레임으로 가장 최근의 센서스에서 나온 데이터들이며, 적어도 96% 이상의 주소가 완전하게 수록된 자료이다. 그룹쿼터 프레임은 가구 단위에 포함되지 않는 하숙 집, 호텔방, 시설 등의 나머지 거처들을 포함한다. 3개의 프레임이 SIPP 표본 주소의 거의 90%를 제공한다. 포함률-개선 프레임은 센서스 당시에는 발견되지 않았지만, 사후 조사에서 발견된 주소(address)를 포함한 프레임을 말한다. 이 프레임에서 추출되는 주소는 매우 적은 수이다. 신축건물 프레임은 신축된 건물들의 프레임을 제공한다. 각 표본 PSU에서 추출 틀에 있는 주소들은 집락으로 그룹화되어 있으며, 이 집락들을 표본으로 추출한 다음 추출된 주소 집락들이 조사에 사용된다.

SIPP의 가중값은 다음과 같은 항목으로 구성된다.

- ① 표본 단위의 추출확률을 나타내는 기본가중값(Base Weight)
- ② 집락 내에서 부차추출에 대한 조정 가중값
- ③ 이주자를 위한 가중값(2차 웨이브 이후)

- ④ 무응답자들에 대한 무응답 조정 가중값
- ⑤ 기지의 모집단 총계로부터 수정을 위한 사후층화 가중값

마. 캐나다의 SLID(Survey of Labor and Income Dynamics)

1) 패널 설계

SLID는 2개의 패널로 이루어져 있는데 1993년 1차 패널이 추출된 후 1996년 2차 패널이 추출됐고 현재는 10개 지역을 포괄하고 있다. 표본에서 인디언 보호구역과 군인, 기관 또는 시설 거주자 등은 제외하였다. 패널 구성은 3년마다 변동되며, 전체적인 패널의 구성은 [부도 1]과 같은 형태로 나타낼 수 있다. 1993년부터 1998년까지를 1차 패널, 1996년부터 2001년까지를 2차 패널로 볼 수 있으며, 이에 따라 1996~1998년에는 2개의 패널이 동시에 존재하는 형태를 가지고 있다. 1999년에 1차 패널이 종료되고 이후 3차 패널이 추가되었다. 2001년 2차 패널이 종료되었으며 2002년 4차 패널이 시작되었다.

[부도 1] SLID의 패널 구성 형태



자료: 손창균(2012). p.31. 재인용.

이처럼 패널이 연동 형태를 이루게 된 것은 횡단면 표본에서는 대표성을 가지게 하고 종단적으로는 패널이 이탈하는 것을 감소시켜 표본을 장기적으로 유지하기 위함이다. 각 표본은 1년에 2번의 조사를 받게 되는데, 첫 번째 조사 때는 노동시장, 두 번째 조사 때는 소득이나 임금에 대한 질문에 응답한다. 소득 관련 조사 때는 통계청의 소득 자료를 사용할 수 있도록 하고 있다.

캐나다 노동력조사(LFS)에서 조사 대상이 된 표본에서 패널의 표본을 추출하는데 이때 층화다단계 추출 방법을 사용한다. 즉 SLID의 패널은 먼저 LFS 중 2개 그룹을 추출하고 최종적으로 1만 5천 가구를 1차 패널로 구성한다. SLID의 표본으로 추출되면 6년간은 패널지속기간으로 인정되어 표본으로 남게 된다.

SLID는 개인뿐 아니라 가구의 특성과도 관련이 있는데, 표본 대상과 거주지를 같이하고 있는 개인은 조사 대상이 된다. 따라서 횡단면 표본은 기존의 조사 대상과 같이 살고 있는 사람으로 구성된다. 이러한 조사 대상을 동거인(cohabitants)이라고 하며, 가중값은 이러한 조사 대상을 바탕으로 횡단면과 종단면이 모두 고려되어야 한다. 종단면 가중값(longitudinal weight)과 횡단면 가중값(cross-sectional weight)은 각각 추출된 표본이 대표성을 가지게 하기 위해 산출된 값이다.

2) 패널 가중값 계산

SLID의 가중값은 우선 종단면 가중값과 횡단면 가중값으로 구분할 수 있다. 초기 종단면 가중값은 가구추출확률의 역수와 같고, 모든 종단면 가구원은 동일한 종단면 가중값을 가진다. 먼저 초기가중값을 정의하고, 웨이브마다 종단면 개인을 그룹화하여 0인 가중값이 생기지 않도록 한

다. 즉 무응답자, 횡단면적으로 응답자의 범위에 속하는 자, 횡단면적으로 범위 밖의 개인 등으로 구분한다. 또한 소득과 근로설문 두 가지 중 적어도 한 가지에 응답한 가구는 응답가구로 취급한다. 만일 두 가지 설문에 모든 가구원이 무응답한 경우에는 무응답가구로 고려한다. 또한 SLID는 응답가구에서 무응답가구원은 응답자로 고려하였다. 응답가구값으로 무응답가구원값을 대체(imputation)하였다. SLID에서 무응답 조정인자는 응답자로 고려된 모든 개인의 가중값에 적용하였으며, 이때 응답가구의 어린이와 조사범위 밖의 개인들은 제외하였다. 이러한 무응답가중값 조정은 사후층화로 수행하였다.

다음으로 무응답 조정을 위해 SLID는 개인의 무응답 조정승수(adjustment factors)를 응답자 그룹(Response Homeogenous Group: RHG)의 응답률의 역수로 정의하였다. 적절한 무응답 승수를 계산하기 위해 무응답을 정의하고, 이에 필요한 설명변수를 선택한 다음 로지스틱 회귀를 이용하여 적절한 응답자 그룹을 형성하여 무응답 승수를 구성하였다. 이러한 과정으로부터 무응답이 조정된 가중값을 정의할 수 있다.

SLID는 횡단면적인 소득의 추정치와 분산을 영향력 관찰치를 이용하여 극단가중값으로 조정하였다. 영향력 관찰치에 대한 조정값은 0과 1 사이의 값으로 계산되며, 2차 패널의 2차 웨이브에서 단지 2개의 개인의 영향력 관찰값이 존재하였다.

다음으로 사후층화를 2개의 패널에 대해 독립적으로 수행하였으며, 최종적으로 사후층화 가중값에 대해 가구와 개인의 정보 보호를 위해 일정 수준의 잡음(noise)가중값을 결합하여 외부 자료로 공개하고 있다.

SLID의 횡단면 가중값은 특정한 기준 연도에서 추정치를 계산하기 위해 2개의 패널이 결합하는 형태의 가중값으로 나타난다. 횡단면적 가중값의 목표모집단은 기준 연도의 12월 31일 현재 10개 province에 거주하는 자

로 보호지역 또는 시설 및 군대시설 거주자는 제외된다. 모든 종단면 개인과 그들과 동거하는 자는 횡단면 표본으로 간주한다. 초기 횡단면 가중값은 무응답이 조정된 종단면 가중값으로 각 시점에서 모집단을 대표한다.

횡단면 가중값 조정의 첫 번째 단계는 무응답이 조정된 종단면 가중값을 배분승수(allocation factor)를 이용하여 2개의 패널 표본을 결합하는 것이다. 또한 횡단면 가중값에 대해 패널 결합 후 중요한 작업의 하나가 패널가중값 배분(weight share)이다. 패널조사에서 표본 가구 내에 이사를 오거나, 가구 내에서 어린이가 새로 태어날 수도 있다. 즉 원표본에 포함되지 않았던 개인들이 신규 표본(동거인)의 일부로 들어오는 경우가 발생한다. 따라서 이들은 추출확률이 없기 때문에 새로운 가중값을 적용할 필요가 발생하게 된다. 이러한 작업이 가중값 배분 작업이다.

SLID는 2개의 패널을 결합하는 방식이므로 1차 패널에 속한 가구의 개인이 2차 패널의 개인에 포함될 가능성이 항상 존재하게 된다. 즉 각기 다른 2개의 횡단면 표본으로 간주할 경우 동일한 개인이 2개의 횡단면 표본에 동시에 속하게 됨으로써 독립적인 패널가중값을 산정하게 되면 한 개인에 대해 서로 다른 2개의 가중값을 고려해야 하는 문제가 발생하기 때문이다. 가중값 배분 과정에서 두 가지의 서로 다른 가중값이 개개인에 대해 계산되는데 하나는 표본가구의 개인들에게 부여되는 개인가중값(individual weight)이고 다른 하나는 표본가구의 모든 가구원에게 동일하게 부여되는 통합가중값(integrated weight)이다.

개인가중값에 대한 분배 과정은 우선 종단면 표본으로 선택된 후 동거인이 이주해온 가구를 구별해야 한다. 해당 가구가 종단면 개인 전체로 구성된 경우에는 가중값 배분 작업이 필요하지 않다. 해당 가구의 모든 구성원에 대해 패널분배승수가 적용된 수정된 횡단면 가중값으로 동일하게 횡단면 가중값을 부과하면 된다.

다음으로 새로운 동거인이 생긴 가구에 대해서는 모든 종단면 가구 구성원들의 무응답이 조정된 횡단면 가중값을 모두 더해야 한다. 그 다음으로 종단면 개인들과 초기에 함께 살았던 동거인의 수를 구한다. 초기에 적어도 1인 이상의 동거인이 있거나 초기에는 없었지만 동거인이 생긴 경우, 모든 구성원의 횡단면 가중값은 종단면 개인들의 가중값을 더하여 가구 내의 종단면 개인들과 동거인들의 수를 더한 값으로 나눈 결과가 된다.

또한 지역 내에서의 이주에 따른 가중값 조정을 수행해야 한다. 이는 SLID의 종단면적 성질과 직접적으로 관련이 있으며, 시간이 지남에 따라 표본에 속한 사람들이 현재의 지역에서 다른 지역으로 이주하게 된다. 이러한 효과를 가중값에 반영할 필요가 있다. 즉 추출확률이 낮은 지역에서 추출확률이 높은 지역으로의 이주로 인한 가중값 변동을 고려해야 분석 과정에서 왜곡을 피할 수 있기 때문이다. 마지막으로 사후층화 조정된 가중값에 동일한 가중값을 갖는 동일한 가구의 개인들의 가중값을 정의한다.

제3절 결론

패널조사는 종단연구 방법으로 시간의 흐름에 따라 변화하는 개인이나 가구의 양태를 통계적으로 분석하는 데 매우 중요한 자료 수집 방법이다. 따라서 개인이 속한 가구의 변화에 대한 지속적인 추적 관리가 필요하며, 이 과정에서 합리적인 추적규칙이 요구된다.

본고는 일반적인 추적규칙과 그에 대한 가중치 조정 과정을 ‘BHPS’, ‘HILDA’, ‘PSID’, ‘SIPP’, ‘SLID’ 등 해외 주요 패널들의 사례를 통해 고찰하였고, 이를 통해 한국복지패널의 추적규칙과 가중치 조정에 대한 가이드라인을 제시하는 데 초점이 있다.

가구 내 가구원의 속성에 따라 어느 시점까지 지속적으로 추적할 것인가에 대한 결정은 패널조사에서 패널 구성의 성격과 맥을 같이하기 때문에 매우 중요한 이슈 중 하나이다. 원가구원에 대한 추적시점과 종결시점의 결정에 따라 가중치 조정 단계가 결정되고 이는 결과적으로 조사자료의 대표성과 관련되며 패널 표본의 규모도 이에 좌우되기 때문이다. 이러한 측면에서 한국복지패널의 패널 추적규칙의 경우 1단계 분가 가구원에 대한 추적까지를 허용하며, 해당 분가가구원이 재분가한 경우 더 이상 추적하지 않는 것으로 파악되었다.

한편 패널가구에 진입하는 신규 가구원의 처리 또한 패널의 추적규칙과 연관되어 패널 표본 규모의 관리에서 중요한 이슈이다. 즉 패널가구에서 새롭게 태어난 신생아의 경우 대부분의 가구 패널에서는 원가구원으로 처리하고 있으며, 따라서 가구가중치에 이를 반영하고 있으나 일부 패널에서는 이를 고려하지 않고 있다. 결과적으로 패널의 특성에 따라 원가구원(또는 원표본)의 성격을 규정하는 규칙이 다르기 때문에 별도로 정해진 규칙은 없는 것으로 파악되었다.

하지만 한국복지패널의 경우 이러한 신생아의 패널화가 유지되고 있으며, 이들은 실제로 응답 대상은 아니지만 지속적으로 관리되는 기록가구원의 성격을 가지게 된다. 본문에서도 언급한 바 있지만 응답가구원과 기록가구원의 성격은 패널 유지 및 추적에 중요한 기준이 되며 이는 가구 및 개인가중치 조정에 반영되기 때문에 명확히 규정될 필요가 있다.

기록가구원은 가구가중치 조정과 개인가중치 조정에 관련되며, 응답가구원은 개인가중치 조정에 관련된다. 한국복지패널의 경우 가구원데이터에서 기록가구원과 응답가구원의 규정이 명확하기 때문에 가중치 처리에는 문제가 없는 것으로 파악되었다.

이상과 같이 알아본 해외 주요 패널의 추적규칙과 가중값 조정 방법의

비교 분석을 통해 향후 한국복지패널의 패널 유지 관리를 위한 가이드라인 구성에 도움이 되길 기대한다.

간행물회원제 안내

▶ 회원에 대한 특전

- 본 연구원이 발행하는 판매용 보고서는 물론 「보건복지포럼」, 「보건사회연구」도 무료로 받아보실 수 있으며 일반 서점에서 구입할 수 없는 비매용 간행물은 실비로 제공합니다.
- 가입기간 중 회비가 인상되는 경우라도 추가 부담이 없습니다.

▶ 회원종류

- 전체간행물회원 : 120,000원
- 보건분야 간행물회원 : 75,000원
- 사회분야 간행물회원 : 75,000원
- 정기간행물회원 : 35,000원

▶ 가입방법

- 홈페이지(www.kihasa.re.kr) - 발간자료 - 간행물구독안내

▶ 문의처

- (30147) 세종특별자치시 시청대로 370 세종국책연구단지 사회정책동 1~5F
간행물 담당자 (Tel: 044-287-8157)

KIHASA 도서 판매처

- | | |
|---|---|
| ■ 한국경제서적(총판) 737-7498 | ■ 교보문고(광화문점) 1544-1900 |
| ■ 영풍문고(종로점) 399-5600 | ■ 서울문고(종로점) 2198-2307 |
| ■ Yes24 http://www.yes24.com | ■ 알라딘 http://www.aladdin.co.kr |