

서구 주요 패널의 조사전략 변화와 시사점

이 상 호*

이제 한국노동패널(이하 KLIPS)이 9차년도 조사로 접어들고 있다. KLIPS는 8차년도 조사에서도 77.6%의 원표본유지율을 기록했을 뿐만 아니라, 각 지역별 성공률 격차의 완화, 직접면접 비율의 향상 등 양적·질적인 측면 모두에서 안정화 단계에 접어들고 있다. 또한 자료 가공단계에서도 체계적인 오류검증 절차들을 구축하고, 매년 학술대회를 개최하여 패널이용의 저변 확대를 위해 노력하고 있다. 그러나 다른 한편으로는 한국의 인구동학과 노동시장 구조의 변화, 컴퓨터 기술의 발전과 국내 조사환경의 변화 등으로 인해 패널조사를 둘러싼 내외적 변화가 급격하게 진행되고 있다. 따라서 향후에는 이러한 변화에 대응하여 노동패널의 조사전략에 있어서도 상당한 수준의 질적 전환이 필요할 것으로 예상된다. 즉, 지금까지의 초점이 주로 시스템 자체를 안정화시키는 것에 있었다면, 향후에는 기존 시스템의 안정성을 훼손하지 않으면서 어떻게 이런 변화를 제대로 착근시킬 수 있을 것인가의 문제가 될 것이다.

흥미로운 점은 서구의 가구패널조사 또한 노동패널과 비슷한 시기에 이러한 도전에 직면했다는 점이다. 대표적인 가구패널인 미국의 PSID, 영국의 BHPS, 독일의 GSOEP, 캐나다의 SLID 등은 짧게는 10여년, 길게는 30년 이상의 기간을 거치면서 나름대로의 전략적 대응을 통해 이런 도전들을 극복해 왔다. 여기서는 이러한 과정 중에서도 표본 구성, 조사도구와 조사방법, 조사현황, 그리고 국가간 네트워크의 형성에 초점을 맞추어 살펴볼 것이다. 마지막으로 이들의 경험을 통해 우리가 얻을 수 있는 시사점은 어떤 것인지 간략하게 짚어볼 것이다.

* 한국노동연구원 데이터센터 노동패널팀 연구원(shlee@kli.re.kr).

I. 미국의 PSID

1. 표 본

PSID가 처음 설계되기 전 미국의 통계국(Bureau of the Census)은 존슨 정부의 ‘빈곤과의 전쟁’(The war on poverty)과 발맞추어 기초자료 수집에 나섰다. 이를 위해 30,000가구를 표집하여 1966년과 67년 두 차례에 걸쳐 SEO(Survey of Economic Opportunity) 조사를 실시하였다. 이듬해인 1968년 통계국은 미시간 대학 SRC(조사연구소)에 의뢰하여 SEO에서 빈곤층 2,000가구를 추출하여 5년 동안 추적조사할 것을 의뢰하였다. 이에 SRC는 빈곤층뿐만 아니라 국가의 대표 표본을 구축하기 위해 추가로 3,000표본을 구성하였으며 그 결과 PSID의 기본축이 되는 5,000여 가구에 이르는 원표본이 구축되었다. 이후 추가표집에 따라 2001년에 표본이 두 부분으로 분리되었다. 첫번째는 원표본(Initial core sample)인 1968년 표집가구 및 분가가구들로서 이들은 2001년 기준으로 7,255가구(조사 성공 6,915가구+비성공 340가구임)이었다. 두번째는 1968년 이후 미국으로 이주한 표본(Immigrant Refresher Sample)으로 792가구가 1999년 조사부터 새롭게 추가되었다.

2. 조사도구와 방법

조사도구와 관련된 변화는 세 기간으로 구분할 수 있다. 첫번째, 1968-1972년 기간 동안은 면접원이 종이로 작성된 설문지에 응답자와 ‘직접대면(face-to-face)’하여 기입하는 PAPI(Paper and Pencil Interviewing) 방식으로 조사가 이루어졌다. 그러다가 1973부터 1992년까지는 PAPI를 유지하되 전화를 이용하여 조사하는 방식으로 바뀌었다. 마지막으로 1993년부터 컴퓨터에 설문지 프로그램을 내장하여 전화로 수집된 정보들을 면접원이 입력하는 CATI(Computer-Assisted Telephone Interviewing) 방식으로 이행했다. CATI와 같은 컴퓨터 지원 면접시스템은 조사원이 응답을 입력하는 순간 곧바로 데이터베이스화되기 때문에 면접시간 및 데이터 가공시간을 단축시킬 수 있다는 장점이 있었다. 또한 입력오류를 줄이고 에러클리닝 시스템을 프로그램안에 내장할 수 있기 때문에 조사의 효율성과 자료의 정확성을 동시에 향상시킬 수 있었다. 그러나 이 과정에서 많은 문제점이 발생하기도 하였다. 우선, CATI를 도입하기 위해 프로그램을 설계하고 테스트하는 등 15개월의 준비기간을 가졌음에도 불구하고 면접시작 시기가 두 달여 가까이 지연되었다. 현장실사기간 중에는 시스템 에러가 발생하여 면접 진행이 상당기간 중단되기도

하였다. 이 밖에도 사전에 예측할 수 없었던 수많은 어려가 발생하였다. 이로 인해 면접원과 응답자간의 신뢰관계가 훼손되어 결과적으로 응답률 저하로 이어지는 것은 물론이고 데이터 편집 및 최종 데이터 공개 등 일련의 모든 과정들도 연기되었다. 그러나 이러한 문제점들은 1994년 이후 안정화되기 시작하였으며 현재는 이행에 성공하였다.

2003년부터는 새로운 CATI 프로그램을 추가적으로 도입하였다. 즉, 주거사항의 변화 및 직업력(work history)의 변화과정을 ‘Event History Calendar(EHC)’기법을 이용하여 기록하였다. 이를 이용하여 면접원이 2년간의 직업력 변화를 각 시점별로 정확하게 입력할 수 있게 되었으며, 보다 정확한 정보수집이 가능해졌다.

3. 조사현황

2003년 조사는 3월부터 9월까지 약 6개월에 걸쳐 이루어졌다. 조사는 SRC 조사팀에 소속된 약 90여명의 면접원들에 의해 96.8% 정도가 CATI로 이루어졌다. 조사결과 ‘1968년 표본’의 경우 직전 조사 대비 96.8%(7,870가구)의 성공률을 보였으며, 이주민 표본은 94.3%의 성공률을 기록했다. 평균 조사시간은 68.7분(접촉시간 및 사후보강 접촉시간 포함할 경우 78.7분)이며, 응답 사례금으로는 55\$가 지급되었다.

II. 영국의 BHPS

1. 표 본

BHPS는 엑세스 대학의 ISER(Institute for Social and Economic Research)에서 실시하는 가구패널조사로 1991년 1차년도가 시작된 이래 현재 13차(2003년) 자료까지 공개되고 있다. 1994년 조사에서는 BHPS조사 응답가구 중 11~15세까지의 청소년을 추출하여 청소년패널(BYP: British Youth Panel)을 시작하였다. 1997~2000년까지는 유럽단위의 가구패널조사인 ECHP에 자료를 제공하기 시작했는데, 그 일환으로 원래 EHCP 중 영국에 할당된 표본들(북아일랜드 포함, 빈곤층 표본)을 BHPS에 통합시켰다. 1999년부터는 스코틀랜드와 웨일즈 지역 표본이 추가되었다. 원래 이들 지역의 표본수가 너무 적다고 판단되어 분석에 필요한 충분한 사례수를 확대하는 것이 첫번째 이유였고, 정부정책의 변화가 미치는 효과를 잉글랜드와 비교하는 것이 두번째 이유였다. 이로써 기존에 400~500가구에 그쳤던 이 지역 표본에 2,000가구가 추가로 표집되었다. 2001년에는 북

아일랜드 표본(약 2,900가구 5,200여명)이 추가되었다. 이로서 BHPS는 2001년부터 영국의 전지역을 포괄하게 되었다.

2. 조사도구와 방법

BHPS는 9차년도(1999년) 조사부터 CAPI시스템을 도입하였다. CAPI로의 성공적인 이행을 위해서 ISER는 실사업체인 NOP와 긴밀한 협력관계의 형성과 명확한 역할분담을 통해 CAPI프로그램과 조사연구팀원들 간의 의사소통 문제를 해결해 나갔다.

도입 첫해에는 가구용 및 개인용 설문에만 CAPI를 사용하고 다른 설문에는 PAPI를 유지하였다. 그러다가 13차조사(2003년)부터는 대리용 설문(장기간 출타중이거나 고령자에게 실시)에도 CAPI를 적용함으로써 이행을 완결지었다. ISER에서는 CAPI로의 이행 과정에서 발생할 수 있는 오류를 최대한 줄이기 위해 신중한 이행과정을 거쳤다. 우선, 가구용과 개인용 설문만을 CAPI로 전환하고, 청년층, 전화조사 등은 기존의 'PAPI' 시스템을 유지했다. 또한 가구용과 개인용 설문에 대해서도 'PAPI' 면접을 함께 실시함으로써 안전장치를 마련해 두었다. 즉, CAPI로 제작된 설문지 구조가 잘못되어 에러가 발생하더라도 이러한 이중구조를 통해 각각의 데이터가 별도로 구축함으로써 정보의 손실을 방지할 수 있었다. 더불어 CAPI로의 성공적인 이행에는 ISER과 NOP의 밀접한 협력관계와 명확한 역할분담이 핵심적인 조건이었다. 다른 한편, 적극적인 협력과 동시에 명확한 역할분담도 중요한 요소였다.

CAPI 시스템에 사용된 프로그램은 기본적으로 SPSS MR에서 개발된 'In2itive'로 이루어졌다. 그러나 북아일랜드 표본에 대해서는 'Blaise' 프로그램으로 수행하였다).

BHPS는 기본적으로 모든 조사가 대면접촉을 원칙으로 하지만, 이것이 불가능할 경우 대리면접을 허용하고 있으며, 3차년도부터는 전화조사도 일부 도입되었다. 또한 개인의 프라이버시를 침해할 수 있는 설문에 대해서는 '자기기입식(self-completed)' 설문도 사용하고 있다.

3. 조사현황

BHPS의 현장실사는 민간 조사업체인 NOP에 위탁하고 있다. 1차년도에 243명의 면접

1) 북아일랜드 표본은 원래 유럽단위의 가구패널조사인 ECHP 표본을 통합시킨 것이다. 그런데 ECHP가 1차년도 조사(1994년)부터 Blaise를 이용한 CAPI시스템을 사용하고 있었기 때문에 BHPS에서도 이를 그대로 유지하고 있는 것으로 보인다. 더구나 현재 BHPS를 운영하는 엑세스 대학의 ISER이 ECHP 전체 데이터센터를 관리하고 있기 때문에 프로그램의 차이에 의한 문제는 그리 크지 않은 듯하다.

원들이 250여개의 표본지역을 포괄하였으며, 2차년도 이후부터는 가급적 동일한 지역에 대해서는 동일한 면접원들이 조사하도록 하였다. 2차년도 조사당시 총 237명이 투입되었는데 이 중에서 35명을 제외한 나머지는 1차년도에도 투입된 면접원이었다. <표 1>에서도 볼 수 있듯이 1~13차까지의 조사기간 동안 약 1/5의 면접원이 한 차수도 빠지지 않고 유지되었음을 알 수 있다. BHPS에는 비표본오차를 관리하기 위한 전략도 마련하고 있다. 2차년도에는 지역효과에서 면접원 효과를 분리하기 위해서 실험도 실시하였다. 즉, 어떤 지역폴 내에서 면접원에게 방문할 가정의 주소를 무작위적으로 할당하였다. 각각의 지역폴은 인근의 2~3개 조사구로 구성되었다. 최종적으로 테스트가 적절한 153개의 조사구를 70개의 지역폴로 구성한 35개 지역폴에서는 면접원이 자신의 조사구 외에 다른 조사구에도 서로 조사할 수 있도록 하였다. 이러한 방식으로 실험한 결과에 따르면 ‘면접원이 동일한 가구를 계속 조사하더라도 거기에 따른 면접원 효과가 존재한다는 증거는 없는 것으로 나타났다(O’Muircheartaigh, Campanelli and Smith, 1999; Campanelli and O’Muircheartaigh, 1999).

면접원 교육은 ISER과 NOP가 공동으로 실시하였는데 기존 면접원의 경우 1일, 신규 면접원의 경우 2일의 프로그램을 별도로 가동시켰다. 9차년도부터는 컴퓨터를 이용하여 CAPI 설문지에 대한 교육을 별도로 실시하였으며, 모든 교육과정에는 ISER 직원들이 참석하였다. 조사는 매년 9월 첫째 주부터 12월 첫째 주까지 이루어지며, 대체로 80% 이상의 조사가 9, 10월에 완료되었다. 그러나 9차년도 조사는 CAPI로 이행하는 과정에서 기술적인 문제로 인해 다음해 1월에서 5월까지 조사가 진행되었다.

응답자 관리에서 특징적인 것은 영국의 경우 현금인센티브를 제공하지 않는 대신 상품권(gift voucher)을 제공한다는 점이다. 1~5차년도 기간 동안은 5파운드 상당의 상품권이었으나, 6차년도 조사부터는 7파운드로 인상되었다. 또한 기존에 응답했던 개인에 대해서는 응답 후가 아닌 응답 전에 조사대상 표본에 대해 이를 우편으로 발송하였다. 신규 응답자나 이전 차수에 응답하지 않은 개인, 청년층조사 응답자들에 대해서도 별도의 상품권을 지급하였다.

<표 1> BHPS의 면접원수

조사 차수	총면접원수	신규 면접원수	1차년도 면접원수	계속조사 면접원 비율(%)
1차년도	243	-	243	100
2차년도	237	35	202	59
3차년도	216	14	181	48
9차년도	212	15	118	25
13차년도	261	25	79	20

III. 독일의 GSOEP

1. 표 본

GSOEP은 주요 패널조사 중에서 가장 다양하고 복잡한 형태의 표본을 구성하고 있다. 1984년 1차조사에는 서독지역만을 표집했으나, 통일 6개월 전인 1990년 6월에 동독지역으로 표본을 확대하였다. 1차년도 조사당시 서독지역 표본은 다시 독일인 표본(표본 A) 4,528가구와 터키, 유고슬라비아, 스페인, 이탈리아인 가구주로 구성된 외국인 표본(표본 B) 1,395가구로 나누어진다. 1990년에 시작된 동독 거주자 표본은 외국인이 제외된 독일인 표본(표본 C)으로만 구성되었다. 1994/1995년도에 시작된 이민자 표본(표본 D)은 표집시기에 따라 D1(236가구)과 D2(295가구)로 나뉘며, 가구원 중에서 1984년 이후 외국에서 서독으로 이주한 가구원이 1명 이상인 가구로 구성된다. 1998년에는 기존의 표본들과는 독립적으로 서독인 표본(표본 A와 유사한 방식)의 추가가 이루어진 표본 E가 추출되었다. 2000년에는 독일인과 이주자 표본을 모두 포함하는 표본 추가(표본 F)가 다시 한번 이루어졌다. 마지막으로, 2002년에는 소득분포의 양 극단에 해당되는 가구들이 지나치게 과소표집되었다는 판단하에 새로운 표본을 구성하기로 하였다. 이 중에서 ‘홈리스’와 같은 극빈계층은 추적조사가 어려울 뿐더러 결과의 신뢰성도 확보하기 어렵다고 판단하여 결국 고소득층에 대해서만 추가표집을 실시하기로 결정하였다. 그리하여 2002년에는 월소득이 7.500DM(3,835EURO) 이상인 가구들에서 ‘표본 G’에 대한 추가표집이 이루어졌으며, 2003년부터는 이 중에서도 4,500EURO 이상인 가구들만을 선별하여 추적하였다. 이러한 지속적인 표본추가 결과 2000년 이후에는 매년 20,000가구 이상의 표본이 조사되고 있다.

<표 2> GSOEP의 표본 특성

표 본	시작연도	가 구	개 인	표 본
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	월소득이 7.500DM(3,835EURO) 이상인 가구

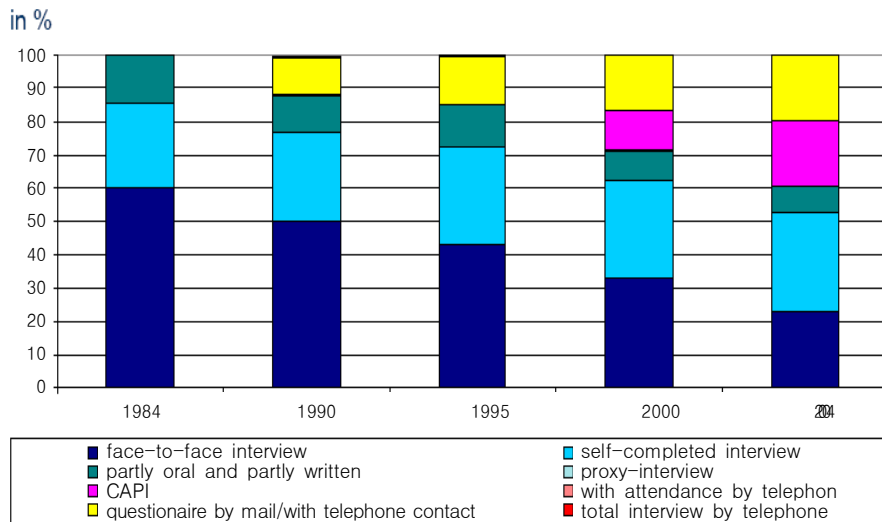
2. 조사도구와 방법

이상과 같은 다양한 표본 구성은 다양한 설문지 설계를 필요로 했다. 따라서 각각의 표본에 따라 상이한 구조의 설문지가 설계되었는데, 1996년 이후로는 표본에 따른 설문지 차이는 없어지고 데이터 제공 또한 통합적으로 이루어졌다. 조사추적은 만 16세 이상의 모든 개인을 추적하는 것이 원칙이다. 각 표본의 1차조사에 해당되는 개인뿐만 아니라, 그 자녀들, 그리고 표본가구에 잠시 진입하였다가 이탈한 개인들까지도 모두 추적하며, 이들에게도 가중치들이 부여된다.

조사원칙에 있어서 GSOEP이 가장 특징적인 것은 모든 가구원 정보들은 본인으로부터 직접 얻는다는 것이다. PSID나 SIPP 등은 가구주로부터 다른 가구원들에 대한 정보를 얻는 형태 등의 대리응답이 사용된다는 점과 비교할 때 매우 차별적이다. 대신 응답자 본인을 통한 정보수집이라는 원칙과는 대조적으로 본격적인 조사이전의 접촉방식에 대해서는 응답자의 상황과 조건에 따라 최대한 유연하게 운영한다.

1998년부터는 CAPI시스템을 단계적으로 도입하기 시작했다. 우선 CAPI 조사대상을 연령이나 하위표본에 따라 선별해서 진행하는 한편, 표본 및 설문지 종류에 따라 PAPI와 CAPI를 혼합하였다. 또한 조사도구로 인해 발생할 수 있는 여러 가지 효과들을 분석하기 위해 응답자 및 면접원들의 반응에 대한 정보들도 수집분석하였다. [그림 1]에서도 볼 수 있듯이, 개인설문지의 약 13%(2005년)가 CAPI로 진행되었으며, 현재까지 체계

[그림 1] GSOEP의 조사방법의 변화(표본 A)



주 : Schupp(2005)에서 인용.

적인 ‘모드효과’²⁾(Mode effect)는 없는 것으로 나타나고 있다.

3. 조사현황

GSOEP에서도 BHPS와 거의 유사한 방식의 면접원 운영원칙을 갖고 있다. 가능한 한 조사지역을 세분하여 ‘구역’(cluster)을 정한 후, 면접원 1인이 각 구역 1개씩을 맡아서 조사한다. 따라서 동일한 면접원을 계속 유지할 경우 조사의 안정성도 높아지지만 대신, 숙련된 면접원이 이탈할 경우 해당 구역의 성공에도 높은 영향을 미친다는 단점이 있다.

응답자 관리를 위해서는 영국의 BHPS와 마찬가지로 현금지급 방식을 직접적으로 사용하지 않는다. 대신 1987년(4차조사) 이후로는 5~10마르크 상당의 선물(서독의 시간당 임금률보다 낮음)을 지급하고 있다. 이 외에도 특정 연도의 경우 응답자의 요청으로 GSOEP의 결과를 설명한 책자를 발송하기도 했다.

이상과 같은 관리시스템하에서 GSOEP은 PSID 다음으로 가장 긴 가구패널 역사를 이루어왔다. GSOEP은 표본구성이 다양하기 때문에 표본유지율 또한 각 표본에 따라 달리 판단할 수밖에 없다. 우선 1984년에 시작되었던 표본A, B를 기준으로 볼 때 5차년도 이전까지는 상대적으로 높은 이탈률을 보였으며, 5~10차년도까지는 약 78% 수준에서 안정된 표본유지율을 기록하다가 11차년도부터 다시 지속적으로 하락하여 21차년도(2004년)까지 62.9%의 표본유지율을 기록하였다. 그러나 이 중에서도 외국인 표본(B)이 서독인 표본(A)보다 조사성공률이 낮으며, 개인성공률은 가구성공률보다 약 7%p가 낮다.

IV. 캐나다의 SLID

1. 표 본

SLID는 캐나다 통계청(Statcan)에서 1993년부터 시작한 종단면 조사이다. SLID의 표본은 캐나다 노동력조사(LFS: Labor Force survey)에서 추출되었기 때문에 표본설계 역시 LFS와 공유한다. LFS 표본은 6개의 독립된 표본으로 구성되는데, 이들 각각을 ‘로테이션집단’(rotation group)이라 하여 매월 1개 그룹씩 교체한다. 따라서 두 조사의 모집단은

2) 종이 설문지에서 컴퓨터 화면을 통한 설문으로 조사형태가 바뀜으로 인해서 응답자가 이질적인 조사로 느끼게 되어 조사에 거부감을 갖게 되는 효과를 일컫는다.

기본적으로 동일하지만 몇 가지 차이가 존재한다. 첫째, SLID의 표본은 모든 연령을 포괄하는데 비해 LFS는 15세 이상의 인구만을 포괄한다. 둘째, SLID의 표본이 모든 연령을 포괄하지만 노동 및 소득에 관련된 조사는 15세 이상만을 대상으로 실시한다. 셋째, 모집단이 포괄하는 지역에 차이가 존재한다. 어쨌든 이렇게 추출된 SLID 표본은 다시 두 개의 하위패널로 나뉜다. 각 패널은 두 LFS rotation 그룹으로 구성되며 약 15,000가구를 포함한다. 한 패널이 6년 동안 조사되며, 새로운 표본이 3년마다 교체된다. 따라서 두 패널들은 항상 조사기간이 중첩된다. 결국 다른 가구패널조사와는 달리 SLID의 원표본은 6년이 지나고 나면 완전교체되는 것이다.

2. 조사도구와 방법

이러한 표본구성으로 인해 표본가구는 6년 동안 최대 12번의 면접을 수행하게 된다. 1월 조사에서는 주로 지난해의 노동시장 경험, 교육 및 가족관계와 관련된 정보들이 수집되며, 5월 조사에서는 소득에 관련된 정보가 수집된다. 무엇보다 응답부담을 줄이기 위해 캐나다 통계청이 갖고 있는 ‘조세정보’를 매칭하며, 약 80%의 응답자에 대해 조세 기록 매칭이 이루어진다. 이들에 대해서는 5월 조사에서 제외된다.

SLID는 첫 조사인 1993년부터 CAI를 이용하여 설문지를 설계하였다. 이때 사용한 프로그램은 오하이오 주립대의 CHRR에서 개발된 CAI와 자체개발한 FOXPRO라는 DB를 공동으로 사용하였다. 그러다가 2000년에 들어서면서 통계청에서 실시하는 모든 조사에 대해 전면적인 CAI 도입을 검토하기 시작했다. SLID는 이러한 시도의 첫번째 사례로 선택되었으며, 소프트웨어도 ‘Blaise’로 결정하였다. 약 1년 동안 설문지 분석, 디자인, 테스트 끝에 최종적으로 2001년 1월 CATI 설문지가 완성되었다. 그러나 이러한 CATI 프로그램이 7개의 지방사무소에서 실행되었을 때 프로그램상의 문제로 실행되는 시간이 지연되는 등의 문제들이 발생했다. 어쨌든 현재는 원만한 조사가 이루어지고 있다. 특히 날짜와 관련된 정보들(고용스펠, 미취업스펠, 퇴직시점 등)은 CAI 기법으로 인해 자료의 정확성이 매우 향상되었다.

3. 조사현황

SLID는 면접원에 대해서도 다양한 유인체계를 갖고 있다. 특히, SLID는 CATI로 조사를 실시하기 때문에 면접원이 조사팀과 괴리되었다는 느낌을 받기 쉽다. 따라서 면접원 뉴스레터를 발송하고, 조사가 끝나면 면접원들에게도 감사편지를 보낸다. 면접원들은 대부분이 LFS조사를 겸임하고 있으며, 가급적 매년 같은 응답자와 접촉하도록 하고 있

다. 응답자 유인체계에 대해서는 현금보상이 응답률을 높이는데 그다지 기여를 하지 못한다고 보고 있다. 대신 응답자에게 조사에 대한 안내책자나 편지, 그리고 자료의 조사 결과를 요약하여 참여를 유도한다.

SLID는 종단면 및 횡단면적 특성을 동시에 갖고 있기 때문에 표본관리 또한 양자를 모두 나누어 실시한다. 조사 또한 1년에 두 번(1월과 5월)씩 이루어지기 때문에 무응답자 또한 두 기간 모두 거부한 개인으로 정의된다. 종단면 표본의 응답률을 살펴보면 <표 3>과 같다.

<표 3> SLID의 종단면 표본(개인) 유지율

	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
패널 1	93.3	89.6	86.5	83.9	82.4	81.5	-	-	-	-	-
패널 2	-	-	-	89.5	86.7	85.2	82.7	78.5	77.4	-	-
패널 3	-	-	-	-	-	-	83.9	83.0	83.0	79.6	76.4
패널 4	-	-	-	-	-	-	-	-	-	81.2	83.2

V. 국가간 패널 네트워크

1. 유럽 가구패널 컨소시엄(CHER)

CHER은 유럽 각국의 가구패널자료를 비교가능한 데이터베이스로 만들기 위해 개발되었다. EU 14개국 이외에도 스위스, 폴란드, 헝가리, 미국 등 총 18개국이 참여하고 있다. 이 중에서 독일, 영국, 룩셈부르크, 벨기에는 자국의 패널자료를 제공하고 나머지 국가들은 유럽단위의 가구패널조사인 ECHP자료를 제공한다. 자료제공 연수는 1990년부터 최대 2000년까지 가능하다. 수록 내용으로는 기본적인 인구학적 특성과 건강, 교육관련 사항, 고용 및 노동시장 활동, 소득과 지출, 주거, 기타 사회적 관계와 관련된 변수들이다. 제공되는 자료의 형태는 SAS, SPSS, Stata, ASCII 등이며, 이 외에도 'MMM-Database'를 운영하여 사회보장과 관련된 제도 정보들(MISSOC), 고용정책과 관련된 제도 변수들(MISEP), 그리고 거시경제적 지표와 관련된 시계열 정보들(Macro-data)을 제공하고 있다.

2. 국가간 표준화 자료(CNEF)

CNEF는 영국의 BHPS(6,000가구 21,000명)와 미국의 PSID(7,000가구 33,000명), 독일의 GSOEP(6,000가구 20,000명), 그리고 캐나다의 SLID(32,000가구 95,000명) 4개국의 가구패널이 핵심 변수들에 대해 비교가능한 표준화된 자료를 제공하기 위해 만들어졌다. 현재 이용가능한 자료로는 1982년부터 2002년이며 향후에도 계속 업데이트될 예정이다. 제공되는 자료의 형태는 SAS, SPSS, Stata, ASCII 등이며, 가구와 개인 고유번호가 있기 때문에 원자료와도 매칭이 가능하다. 주로 가구소득에 초점을 맞추어 제공하며, 건강, 교육, 주거, 이동, 고용특성에 대한 정보들도 제공된다.

<표 4> 주요 가구패널조사의 조사시스템 비교

	PSID	GSOEP	BHPS	SLID	KLIPS
	미국	독일	영국	캐나다	한국
최초 표본수	4,800가구	5,900가구	5,500가구	각 패널당 약 15,000가구	5,000가구
시작연도	1968	1984	1991	1993	1998
조사기관	SRC	DIW	엑서스 대학 ISER	캐나다 통계청 Statcan	한국노동연구원
현금보상	55\$ (2003년)	로터리 티켓 (5-10DM상당)	7파운드 가량의 쿠폰	현금보상없음	3만원
실사기관	자체조사	TNS Infratest (민간업체)	NOP (민간업체)	통계청지역사무소	HRC (민간업체)
조사주기	격년(1997 이후)	매년	매년	매년	매년
조사대상	가구주	모든 가구원	모든 가구원	가구주	모든 가구원
표본지속주기	제한없음	제한없음	제한없음	6년마다 교체	제한없음
저소득층 가중표집	예	아니오	예	아니오	아니오
인종, 이민자 표본	라틴	외국인, 이민자 (구)동독 지역	북아일랜드, 스코틀랜드, 웨일즈	없음	없음
청(소)년부가조사	있음	있음	있음	있음	3차조사 실시 9차조사 예정
일자리정보기준시점	지난달	지난달	지난해	지난해	현재
조사추적 원칙	원표본	모두	원표본	모두	원표본
정기적인 표본추가	없음	없음	있음	있음	없음
조사도구	CATI	PAPI, CAPI	CAPI	CATI	PAPI
CAI 도입시점	1994년	1998년	1999년	2001년	2007년 예정
CAI프로그램명	Blaise	In2iview	In2iview(Blaise)	Blaise	Blaise 예정
가중치 형태	횡단면/종단면	횡단면/종단면	횡단면/종단면	횡단면/종단면	횡단면/종단면
데이터 추출도구	웹기반 'data center'	CD롬, 웹기반 'SOEPINFO'	없음	CD롬, 윈도우기반 'SLID data Retrieval System'	CD롬
온라인 문헌검색	웹페이지	자체개발S/W 'SOEPLIT'	웹페이지	웹페이지	웹페이지

VI. 노동패널에 주는 시사점

지금까지 살펴보았듯이, 서구의 주요 패널들은 표본, 조사도구 및 방법, 그리고 조사 내용 측면에서 다양한 변화를 거쳐왔다. 특히 1990년대 후반을 기점으로 컴퓨터 지원 면접 시스템을 도입하는 등 새로운 조사체계로의 이행과정이 전개되었다(표 4 참조). 더불어 내적으로는 조사방법론 및 패널분석방법론과 관련된 이론적·경험적 역량을 축적하고, 외적으로는 국가간 네트워크를 구축하여 정보교류 및 시너지 효과도 노리고 있다. KLIPS도 향후에는 CAPI 시스템으로의 이행, 데이터의 종단면 변수들의 표준화, 다양한 축소(맞춤)형 자료의 제공, 조사 및 분석기법에 대한 연구 등과 같이 조사전략 전반에 걸친 어려운 과제들을 풀어내야 한다. 이상으로부터 몇 가지 시사점을 살펴보면 다음과 같다.

첫째, 서구 패널들은 자료의 대표성을 확보하기 위해 각국의 인구동학 및 조사특성에 따라 다양한 형태의 표본전략을 사용하고 있는데, 그 이유는 크게 두 가지로 이루어진다. ① 독일 통일이나 유럽의 EU통합, 그리고 미국의 이민유입 등과 같이 특정 지역이나 인종 등으로 인한 인구구성 변화가 두드러질 때이다. ② 패널조사가 반복됨에 따라 자연스럽게 감소하는 표본의 감소가 일정 수준을 넘어설 때이다. ①의 경우 해당 특성을 가진 표본만을 추가하며, ②의 경우 전체 표본을 추가하는 형식으로 이루어진다. 한국의 경우 유럽과 같은 수준의 이민에 의한 인구유입 및 유출이 낮으며, 최근 도농간 이동보다는 도시간 이동이 주요하게 나타나긴 하지만, 분명 10여년 전과는 다른 형태의 인구 분포가 형성되었을 가능성을 배제할 수 없다. 그간 KLIPS는 조사 설계당시 농촌지역을 제외한 도시지역 표본이라는 점이 한계점으로 지적되곤 했던 점을 감안한다면 새로운 표본구성에 대한 필요성 그 자체는 타당하게 보인다. 그럼에도 불구하고, 현재의 표본유지율 자체가 안정화되어 있기 때문에 조사체계의 다른 변화들을 함께 고려하여 도입 형태와 시기를 결정해야 할 것이다³⁾.

둘째, 조사도구 측면에서는 CAI 도입이 전반적인 흐름으로 자리잡고 있다. 그러나 서구 여러 국가들의 패널 경험들이 공통적으로 지적하는 바는 CAI 도입이 설문지 디자인 뿐만 아니라 현장실사 및 데이터 관리 등 조사체계 전반의 변화를 수반하는 것이기 때문에 오랜기간의 준비와 테스트를 거쳐야 한다는 것이다. 더구나 자료수집의 정확성과

3) 노동패널팀에서는 이미 표본추가 방법 및 형태에 대해 구체적인 검토를 한 바 있다. 자세한 내용은 김재광(2003) 참조.

조사프로시저의 단축을 위해 CAI를 도입하더라도 새로운 형태의 오류가 발생가능하며, 조사준비 단계의 프로시저 역시 비약적으로 증가할 수 있다⁴⁾. 따라서 노동패널에서도 CAI 도입의 필요성은 인정되나 도입시기와 절차, 방법에 대해서는 철저한 준비기간을 가져야 한다.

셋째, 자료의 국제화 작업도 중요한 과제이다. 앞서 언급했듯이 PSID나 BHPS 등은 자신들의 정보를 서로 공유하고 이로부터 시너지를 발생시키기 위해 다양한 형태의 국가간 네트워크를 구축하고 있다. 노동패널 역시 이런 흐름에 참여함으로써 국제적인 위상도 높이는 한편, 이들이 가진 노하우와 경험들을 획득할 수 있을 것이다. 현재 노동패널 팀에서는 데이터와 코드북, 유저가이드 등의 영문화 작업을 단계적으로 진행하고 있다. 또한 국제화를 위해 필요한 자료의 표준화와 같은 여러 가지 기초작업들도 구상 및 추진중에 있다.

넷째, 면접원 및 응답자 관리전략에 있어서도 몇 가지 개선점이 보인다. 주요 패널들이 조사응답률 및 자료의 정확성을 높이기 위해 다양한 면접원 및 응답자 유인체계를 개발하고 있다. 특히 면접원 효과나 현금보상의 효과 등을 엄밀하게 테스트하기 위해 엄밀한 기초하에 실험을 설계하여 분석한다는 것은 우리가 배워야 할 부분이라고 본다. 노동패널의 경우 4차년도부터 가구당 3만원의 현금보상을 실시하고 있지만 이것이 조사에 어떤 효과를 미쳤는지는 확실하게 알 수 없다. 따라서 응답자의 태도나 면접원이 조사에 미치는 효과 등 ‘비표본 오차’(Non-sampling error)를 분석할 수 있는 체계 및 기초 자료의 수집에 대해서도 고민해야 할 것이다.

다섯째, 조사결과에 대한 분석을 보다 깊이 있게 진행할 필요가 있다. 현재 대부분의 가구패널에서는 조사성공률로 가구뿐만 아니라 개인단위 성공률도 상세하게 다루고 있다. 이는 패널조사에서 표본가구와 비표본가구의 정의가 모호하기 때문에 궁극적인 표본단위는 개인이 될 수밖에 없다는 점에 기인한다⁵⁾. KLIPS 역시도 가중치의 부여나 조사추적단위 등 실질적인 면에서는 개인단위도 중요하게 다루고 있다. 따라서 향후에는 개인단위의 성공률을 포함하는 보다 다양한 차원에서의 표본관리와 분석에 신경을 써야 할 것으로 여겨진다.

4) 각국 패널들이 CAI 이행을 거치면서 겪었던 구체적인 과정에 대한 보고서들은 거의 없다. 다만 군데군데에서 이런 과정이 얼마나 복잡하고 힘들었는지를 알 수 있는 대목들이 있다. 예컨대, 던컨(Duncan, 1999)은 CATI로의 이행과정을 ‘끔찍한 악몽’으로 묘사하면서 자신은 이러한 기법의 도입에 관한 한 ‘러다이트’(Luddite)가 되겠다고 공언한 바 있다. BHPS의 보고서에서도 역시 비슷한 표현으로 이행과정을 묘사한다.

5) 예컨대, 4인의 표본가구원으로 구성된 가구에서 2명이 분가할 경우, 어떤 가구가 원표본이고 어떤 가구가 분가가구인지 구분한다는 것은 매우 어려우며, 분가가구라고 해서 조사에 소홀히 해야 할 이유는 전혀 존재하지 않는다.

여섯째, 각국의 패널조사기관들은 조사설계와 자료수집 및 제공뿐만 아니라 패널조사 및 분석방법 그 자체를 개선시키기 위한 체계적인 연구·교육 단위를 갖고 있다는 점에 주목할 필요가 있다. 예컨대, BHPS의 기금을 제공하는 ESRC에서는 다른 대학에서 패널 관련 코스를 개설하고 운영하는 데에도 자금을 지원하고 있다. 노동패널 역시 워킹페이퍼 시리즈 발간, 학술대회 개최, 노동패널 자료설명회 등을 통해 연구 및 교육사업을 하고 있지만 이들 국가의 수준에는 훨씬 못미치는 수준이다. 더구나, 국내에서도 최근 패널조사에 대한 수요가 급증하고 있기 때문에 패널조사 방법 및 분석과 관련된 보다 조직적인 이론적·실증적 성과를 도출함으로써 국내 패널조사 및 연구의 수준향상에 기여할 수 있을 것이다.

그러나 이러한 변화들을 논하기에 앞서 패널조사의 성공을 위해서는 무엇보다 탄탄한 내적 안정성에 흔들림이 없어야 함을 유념할 필요가 있다. 저명한 경제학자이자, PSID에 수십년간 몸담았던 던컨(Duncan)도 ‘The PSID and Me’라는 자전적인 에세이에서 PSID 성공의 가장 중요한 요인으로 조사팀원들의 역할을 꼽았다. 이들의 ‘제도적인 기억’(Institutional Memory)이 암묵적인 지식으로 누적되어 조사차수간에 발생할 수 있는 문제들을 일관되게 극복할 수 있었다는 것이다. KLIPS 역시 조사팀과 실사업체 연구진, 면접원, 응답자들간의 보이지 않는 관계들이 톱니바퀴처럼 맞물려 돌아가고 있기 때문에 이러한 교훈은 마찬가지로 적용될 수 있다. 따라서 이상에서 언급한 변화들 역시 이러한 기존 시스템에 잘 녹아들면서 새로운 호순환을 만들 수 있는 방향으로 이루어져야 할 것이다. **KL**

<참고문헌>

강석훈(1997), 「유럽의 패널조사 현황과 시사점: 가구패널조사와 국제비교학적인 관점을 중심으로」, 한국노동패널 워킹페이퍼 시리즈.

금재호(1998), 「캐나다의 노동패널조사」, 한국노동패널 워킹페이퍼 시리즈.

김재광(2003), 「한국노동패널 오차분석 및 표본추가」, 한국노동패널 워킹페이퍼 시리즈.

신동균(1998), 「미국 패널데이터의 현황과 시사점」, 한국노동패널 워킹페이퍼 시리즈.

Armstrong and House(2005), “Data Quality in the 2003 Survey of Labour and Income Dynamics(SLID)”, Statistics Canada, Income Research Paper Series.

Banks, R, and Laurie, H.(1999), “From PAPI To CAPI while staying Happy: The case of the British Household Panel Survey”, ASC’99 Third International Conference, University

- of Edinburgh. September 22:24, 1999.
- Chan, Richard.(2001), “Survey of Labour and Income Dynamics(SLID) Project at Statistics Canada”, Statistics Canada.
- Dibbs and Hale(1993), “Questionnaire Design In a Paperless Society, SLID Research Paper Series”, Catalogue No.93-08.
- Duncan, Greg J.(1999), “The PSID and Me”, working paper. PSID 홈페이지(<http://psidonline.isr.umich.edu/>).
- EPUNet(2005), “Panel Micro-Databases for Socio-Economic Research in Europe: ECHP, CHER, CNEF, & EPUNet”, Seminar Presentation, EPUNet 홈페이지(<http://epunet.essex.ac.uk>).
- Haisken-DeNew and Frick ed.(2005), *Desktop Companion to the German Socio-Economic Panel*.
- Haisken-DeNew, DIW, and IZA(2001), “A Hitchhiker’s Guide to the World’s Household Panel Ddata Sets”, *The Australian Economic Review*, 34(3), pp.356~366.
- Hill, Martha S.(1992), *The Panel Study of Income Dynamics User’s Guide*, SAGE Publications.
- Kroh and Spieb(2005), “Documentation of Sample Sizes and Panel Attrition in the German Socio Economic Panel(SOEP)”, DIW Data Documentation.
- Lynn, Peter et al(2006), “Quality Profile: British Household Panel Survey”, Institute for Social and Economic Research, University of Essex, BHPS 홈페이지(<http://www.iser.essex.ac.uk/ulsc/bhps>)
- Michaud and Saint-Pierre(1994), “SLID Following Rules: Who to Trace and Who to Interview”, SLID Research Paper Series, Catalogue No.94-01.
- Schupp(2005), “Innovations in survey instruments and survey methodology of the SOEP”, ISER Seminar Presentation.
- Schupp and Wagner(2002), “Maintenance of and Innovation in Long-term Panel Studies: The Case of the German Socio-Economic Panel(GSOEP)”, Discussion Papers, DIW.
- Statcan(2003), “User’s Guide for Cross-sectional Public-Use Microdata File-Survey of Labour and Income Dynamics”, Reference year 2000, Statistics Canada, Income Research Paper Series.
- Webber, Maryanne(1994), “The Survey of Labour and Income Dynamics: Lessons Learned in Testing”, SLID Research Paper Series, Catalogue No.94-07.

◆ 약어 정리

BHPS: British Household Panel Survey

CAI: Computer Assisted Interview

CAPI: Computer Assisted Personal Interview

CATI: Computer Assisted Telephone Interview

CHER: Consortium of Household Panels for European Socio-Economic Research

CHRR: 미국 오하이오 주립대의 인적자원연구센터(Center for Human Resource Research)

DIW: Deutsch Institut fur Wirtschaftsforschung

ESRC: Economic and Social Research Council

GSOEP: German Socio-Economic Panel

HRC: 한국리서치(Hankook Research)

ISER: 영국 Essex대학의 사회경제연구소(Institute for Social and Economic Research)

KLI: 한국노동연구원(Korea Labor Institute)

KLIPS: 한국노동패널(Korean Labor and Income Panel Study)

PAPI: Paper and Pencil Interview

PSID: Panel Study of Income Dynamics

SLID: Survey of Labor and Income Dynamics

SRC: 미국 미시간대학의 사회조사센터(Survey Research Center)

StatCan: Statistics Canada