

농가소득의 동태적 변화요인

- 농가경제조사 자료의 패널화와 그 분석 -

2008. 12.

방태경·강석훈



통계개발원

< 목 차 >

제1장 서론	1
제2장 패널통계의 특징과 현황	4
1. 패널통계의 의미	4
2. 패널통계의 유형	9
3. 패널통계의 현황	10
제3장 패널자료 분석방법론	90
1. 분산분석	90
2. 정태 선형분석	92
3. 동태 선형분석	99
4. 패널 분석의 예	102
제4장 농가경제조사 자료의 사후 패널화	109
1. 농가경제조사의 개요	109
2. 사후 패널화	120
3. 표본유지 및 변동분석	125
4. 중단면 가중치 작성	127
제5장 농가경제조사 패널화 자료를 이용한 분석결과	139
1. 소득이동 분석	139
2. 자산, 부채이동 분석	147
3. 전·겸업별 농가이동	153
4. 농업소득 및 농업순생산 결정요인 분석	155
제6장 결 론	162
1. 시사점	162
2. 한계와 향후 연구과제	164

참고문헌	165
<부록 1> 영국의 노동력조사 패널화 예시	166
<부록 2> 농업소득 및 농업순생산 결정요인분석에서 사용한 변수의 기초통계량	168
<부록 3> 외국의 패널조사 현황(자료: PSID 웹 사이트)	172

<표 차례>

<표 1- 1> 정부출연연구기관의 패널사업 실시현황	2
<표 2- 1> 국내 패널조사	13
<표 2- 2> 패널 설계유형	16
<표 2- 3> 생애주기별 분류	17
<표 2- 4> 조사목적 및 정책 이슈	22
<표 2- 5> 표본추출	27
<표 2- 6> 현장조사 진행	29
<표 2- 7> 응답자 관리방법	31
<표 2- 8> 표본유지율 및 항상응답률(always in)	34
<표 2- 9> 패널조사의 기여점 및 정부정책 활용사례	36
<표 2-10> 자료활용 및 제공	38
<표 2-11> 패널조사 예산	40
<표 2-12> 패널별 인력현황	42
<표 2-13> 조사방법의 변화	48
<표 2-14> PSID의 핵심조사항목	50
<표 2-15> 추가된 주제	52
<표 2-16> PSID 조직도	54
<표 2-17> BHPS의 주요주제	61
<표 2-18> 개인용 설문지의 내용	65
<표 2-19> 조사 연도별 특별 주제	72
<표 2-20> GSOEP 하위표본	74
<표 2-21> 설문지의 내용	80
<표 2-22> PACO 데이터베이스 구성내역	87
<표 3- 1> long form 자료	102
<표 3- 2> wide form 자료	103
<표 4- 1> 조사대상 비교표	110
<표 4- 2> 일계부와 원부 조사방법	110

<표 4-3> 모집단(농가수) 변화추이	112
<표 4-4> 표본추출틀의 층화기준	113
<표 4-5> 농가소득 산출과정	115
<표 4-6> 가구명부 파일의 구조	121
<표 4-7> 가구원 명부파일의 구조	122
<표 4-8> 5년간(2003~2005) 패널화 농가교체 현황	123
<표 4-9> 교체사유별 교체 현황	124
<표 4-10> 가구원 명부 패널화 작업 결과	124
<표 4-11> 기본적인 응답률	125
<표 4-12> 주요 가구패널조사에서의 표본유지율	126
<표 4-13> 경제활동인구조사와 도시가계조사에서의 표본유지율 ..	126
<표 4-14> 응답률 추정모형	129
<표 4-15> 비율조정변수의 기초통계량	132
<표 4-16> 가중치의 기초 통계량	132
<표 4-17> 가중치를 이용한 통계량 계산결과	134
<표 4-18> 가중치 사용 전후의 변화분석	137
<표 5-1> 5분위별 농가소득 및 농업소득 기초 통계량	139
<표 5-2> 농가소득이동 분석	142
<표 5-3> 농업소득이동 분석	145
<표 5-4> 5분위별 농가자산 및 농가부채의 기초통계량	147
<표 5-5> 자산분위의 이동	149
<표 5-6> 부채분위의 이동	151
<표 5-7> 전·겸업별 농가 이동	154
<표 5-8> 회귀분석 결과	158
<부표-1> 탈락확률이 높은 무응답자의 특성	166
<부표-2> 초기 가중치를 이용한 분석결과	167

농가소득의 동태적 변화요인

- 농가경제조사 자료의 패널화와 그 분석 -

방태경
강석훈*

제1장 서론

최근 가계 및 기업과 관련된 분석에서 패널자료를 이용하는 경우가 급증하고 있다. 이러한 연구수요에 부응하기 위해 국내외에서 다양한 종류의 패널자료가 작성되고 있으며, 특히 한국의 경우 패널자료의 수집이 급격하게 증가하고 있다. 다음의 <표 1-1>은 현재 실시되고 있는 대표적인 정부출연 연구기관의 패널조사를 열거한 내용이다. <표 1-1>에 따르면 정부출연 연구기관의 패널조사에서 사용하는 예산이 2007년에 112억 원에 달하고 있다.

패널자료를 이용하는 경우 일반적인 반복횡단면자료로는 분석하기 어려운 다양한 동태적 분석이 가능하다는 장점이 있으나, 새로운 패널조사를 실시하는 방안은 횡단면조사를 실시하는 방안에 비해 비용이 많이 소요된다는 문제점이 있을 수 있다. 패널조사에서 비용이 추가되는 원인은 패널구성원이 주거지를 이동하게 되는 경우 이를 추적하여야 한다는 점, 이동하지 않은 표본의 경우에도 평상시에 패널유지관리를 위해 비용이 소요된다는 점과 패널자료의 클리닝(cleaning) 과정에서 오랜 시간과 비용이 소요된다는 점 등을 들 수 있다.

* 외부 공동연구자 : 성신여자대학교 경제학과 교수(shkang@sungshin.ac.kr)

2 농가소득의 동태적 변화요인

〈표 1-1〉 정부출연연구기관의 패널사업 실시현황(단위 : 백만원)

개시연도	주관기관	조사명	사업기간	'07년 예산
1998	한국노동연구원	▪ 한국노동패널	1998~계속	850
2001	한국고용정보원	▪ 청년패널	2001~계속	779
2002	한국노동연구원	▪ 사업체패널	2002~계속	500
2003	한국청소년정책연구원	▪ 청소년패널	2003~2008	439
	한국직업능력개발원	▪ 한국교육고용패널	2004~2015	800
2004	한국교육개발원	▪ 한국교육종단연구	2004~2023	250
	한국직업능력개발원	▪ 인적자본기업패널	2004~계속	450
2005	한국노동연구원	▪ 고령화연구패널	2006~계속	769
	국민연금관리공단	▪ 노후보장패널	20052015	700
2006	한국고용정보원	▪ 대졸자직업이동패널	2006~계속	1,761
	한국여성정책연구원	▪ 여성가족패널	2006~2010	770
	한국보건사회연구원	▪ 복지패널	2006~계속	1,452
2007	한국여성정책연구원	▪ 여성인력패널	2007~계속	238
	한국장애인고용촉진공단	▪ 장애인고용패널	2007~계속	570
	육아정책개발센터	▪ 아동패널	2007~2019	50
	한국보건사회연구원	▪ 의료패널	2007~계속	500
	한국보건사회연구원	▪ 인구패널	2007~계속	280
합 계				11,158

자료 : 김용성(2007), 이희길 (2008)

한편, 패널조사 방법이 널리 활용되고 이를 이용한 다양한 분석결과가 제시됨에 따라¹⁾ 통계청에서도 패널자료를 만들어야 한다는 의견이 제시되고 있는 바, 비용측면과 인력측면 그리고 기술적인 측면에서 통계청에서 새로운 패널자료를 만든다는 방안은 적절하지 않다는 의견도 없지 않다.

이와 같은 점을 감안할 때 먼저 통계청이 실시한 다양한 반복횡단면 자료를 패널자료로 연결하여 사용하는 방안은 패널자료를 새로 수집하는

1) 실제로 국내 국책연구기관에서 작성하고 있는 모든 자료는 일반인들에게 공개되며, 대부분의 패널조사들은 매년 학술대회를 통해 패널자료를 이용한 연구결과들을 발표하는 세미나를 개최하고 있다.

것과 유사한 효과를 가질 수 있다는 점에서 이를 먼저 검토할 필요가 있다. 통계청에서 국가예산으로 작성한 자료를 약간의 추가적인 노력으로 가공하여 정책적 시사점을 가지고 있는 새로운 데이터를 생성해 낼 수 있다면 이는 국가예산의 효율적인 사용이라는 측면에서 매우 바람직한 일이라고 할 수 있을 것이다.

이에 본 연구에서는 통계청에서 가지고 있는 기존의 반복횡단조사를 패널자료로 구축하는 방안을 연구하고자 한다. 통계청에서 동일한 대상을 상대로 특정 주제에 대해서 반복적으로 측정하는 반복횡단조사는 비록 표본을 추적조사하지는 않는다는 점에서 전통적인 의미의 패널조사라고 할 수 없으나 매년 상당한 정도의 동일 표본이 조사에 응한다면 동 자료를 연결하는 경우 패널자료의 속성을 가진 데이터를 작성할 수 있게 된다. 이미 미국, 영국, 캐나다, 뉴질랜드 통계청에서 반복횡단조사를 패널자료로 구축하여 제공하고 있다.²⁾

본 연구는 농가경제조사를 이용하여 통계청의 반복횡단면 자료를 패널화하는 방법을 연구한다. 농가경제조사를 선택한 이유는 농가경제조사가 기본적으로는 농가가 이사를 가거나 농업을 그만 두지 않는 한 5년간(표본교체 주기)은 동일 농가를 조사한다는 점에서 패널화하였을 때 웨이브간 표본손실이 가장 적은 조사 중의 하나로 판단되기 때문이다. 또한 FTA 등 경제개방이 확대되면서 농업부분이 상당한 경제적 타격을 받을 것으로 예상되고 있는 상황에서 농가 경제상태의 변화를 동태적으로 파악하는 일은 매우 의미 있는 작업이기 때문이기도 하다. 이미 정부에서는 FTA에 대비하여 다양한 농업경쟁력 제고 정책을 실시하고 있는 바, 이러한 정책들이 구체적으로 농가에 어떠한 영향을 미치고 있는 지를 파악하는 일도 매우 중요한 일이다.

이와 같은 연구목적 이외에도 국내에 다양한 패널데이터가 작성되고 있지만, 그 현황을 전체적으로 종합한 자료가 별로 없다는 점에 착안하여 본격적인 농가경제조사의 패널화 작업결과를 제시하기 이전에 국내 패널조사의 현황을 대하여 먼저 살펴본다³⁾. 또한 패널조사의 분석예를

2) <부록 1>에서는 영국의 노동력조사(Labor Force Survey)를 패널화하는 경우를 예시하고 있다.

3) 국내 패널조사의 현황에 대해서는 이희길 외(2008)의 연구와 백희중 외(2004)의 연구가 있다.

4. 농가소득의 동태적 변화요인

추가함으로써 패널조사의 활용도를 높이는데 기여하고자 한다.

본고의 순서는 다음과 같다. 먼저 2장에서는 패널통계의 특징과 현황에 대하여 살펴보고, 3장에서 패널데이터를 이용하는 분석하는 방법을 이론적으로 소개한 다음 간단한 실증분석 예를 제시한다. 4장에서 농가경제조사의 사후 패널화 과정, 결과 그리고 종단가중치 작성 방법을 소개한다. 5장에서는 농가경제조사 패널화 자료를 이용하여 분석할 수 있는 주제들을 예시하고 실증분석결과를 제시한다. 6장에서는 본 연구의 시사점과 한계 그리고 향후 연구과제 등에 대하여 논의한다.

제 2 장 패널통계의 특징과 현황⁴⁾

1. 패널통계의 의의 : 정의, 특징, 장단점

가. 정의

패널조사(panel survey)는 동일한 표본을 대상으로 동일한 질문을 시차를 두고 반복 추적하는 조사이다. 따라서 그 표본이 포함하는 개개 조사단위에 대해 복수의 조사 결과를 제공한다. 종단조사(longitudinal survey)⁵⁾로도 불리는 패널조사는 조사대상이 개인, 가구, 기업체 등이 될 수 있고, 이들 조사단위를 지속적으로 추적하는 조사이다.

통계학이나 계량경제학에서는 이러한 패널자료를 이차원 자료(two-dimensional data)로 일컫는다(Hsiao, 2005). 대체로 데이터는 차원의 수에 따라 분류되는데, 단일 현상에 대해 복수의 시점에서 관찰을 한 자료가 시계열(time series)로 일컬어진다. 시계열 자료에서는 자료의 값(values)과 순서(ordering)가 모두 의미를 가진다. 반면에 단일 시점에서 복수의 현상에 대한 관찰을 포함하는 자료는 횡단면(cross-sectional)이라

4) 이희길 외(2008) 통계개발원, '국내패널조사의 현황 분석'과 백화종 외(2004) 한국보건사회연구원, '빈곤 및 공공부조 패널 데이터 구축방안'에서 발췌

5) 종단조사는 동일한 현상을 긴 기간 걸쳐 추적하여 자료를 수집한다는 개념으로 넓은 의미를 갖는다. 이는 ① 패널조사(panel survey), ② 회고적인조사(retrospective surveys), ③ 행정자료(조사)(administrative data)로 구분되며, 일반적으로 종단조사를 패널조사라고 혼용하여 사용한다.

한다. 횡단면 자료에서 자료의 값은 의미를 지니지만 자료의 순서는 그렇지 않다. 이와는 달리 복수의 현상에 대해 복수의 시점에서 관찰을 포함하는 자료가 패널자료이다. 예를 들면, 형제나 가족과 같은 동일 표본 집단에 대해 여러 시점의 관찰을 통해 얻은 자료가 패널자료이다. 시계열이나 횡단면 자료는 일차원만을 가지지만, 패널 자료는 이차원 자료이다(Wikipedia, the free encyclopedia).

나. 특 징

횡단조사(cross-section survey)는 한 시점에서의 많은 정보를 제공해 주기는 하지만 변화의 형태를 유발하는 원인의 파악이나 행동의 변화양상을 보여주지는 못한다. 즉 변수들간의 상관관계는 파악할 수 있지만 그것이 인과관계로 추론할 수 없다. 인과관계는 두 변수 간에 시간적인 간격이 전제되어야 추론이 가능하다. 그러나 종단조사(longitudinal survey)는 일정한 대상집단에 대하여 지속적으로 관찰을 수행하여 형태의 시간적 변화를 밝혀내는데 유용 즉 변수들간에 원인변수를 밝혀낼 때 충분히 유용한 자료로 활용될 수 있다.

패널조사(panel survey)는 동일한 표본에 대해 주기적으로 조사를 반복하여 실시하는 조사로 시간이라는 차원이 추가되어 명시적으로 시간을 고려한 역동적 분석이 가능하다. 이에 패널조사는 시간의 경과에 따른 개별 조사단위의 변화양상과 실태 파악의 수준으로도 그 가치가 높으며 횡단적 조사로 불가능한 종단적 과정에 대한 다양한 분석의 기초 자료가 된다.

패널조사와 코호트와의 개념차이가 있는데, 코호트의 어원은 라틴어로 'cohors'로 'co'는 '뜰'을, 'hors'는 훈련받은 사람들로, 즉, 같은 뜰에서 훈련받는 사람들을 뜻한다고 볼 수 있다. 이것이 확장으로서 친구, 동지, 또래, 동료라는 의미, 혹은 더 나아가서 모든 인간의 집합적 무리를 가리키는 것이다(김응력 2001), 이들을 인구통계학자들이 전문용어로 사용하기 시작하면서 코호트를 '일정한 시기에 있는 인생에 있어 동일하고 중대한 사상을 체험한 사람들의 집합'이라고 정의, 즉 코호트라는 용어는 세대간 차이의 연구를 위해 사용되었던 '세대(generation)'를 학문적

6 농가소득의 동태적 변화요인

분석법으로 인구통계학자들이 사용하면서 시작, 코호트는 출생코호트(birth cohort)를 말하겠지만, 결혼코호트, 교육코호트, 과부, 퇴직, 미혼 등의 코호트가 존재할 수 있다. 이 경우 세대는 출생코호트의 하나로 이해될 수 있지만, 엄밀히 말해서 세대는 부모보다 자녀세대간의 관계라는 의미를 함축하고 있는 것이어서 세대보다 구체적인 의미이다.

패널조사의 특징은 사회현상을 미시적(micro analysis), 동태적인(dynamic analysis)인 분석을 하기 위한 조사이다. 즉 시계열분석은 사회경제내에 전체를 분석하기 때문에 가구·사업체에 대한 변화를 포착하기 어려울 뿐만 아니라, 가구·사업체의 구조적인 변화를 감지해 내기가 쉽지 않다. 횡단면 자료는 한 시점에서 사회경제내의 개별 구성에 대한 정보를 기록한 자료로써 직접적으로 사회의 개별 구성원의 경제 활동을 분석은 가능하지만, 이 자료를 통하여 전체 사회경제의 총체적인 구조나 변화의 형태를 분석하기는 어려운 것이 사실이다. 종단면자료는 개별 구성원의 한정된 자료와 개별 구성원의 형태에 영향을 미치리라고 예상되는 구성원 간에 유기적으로 결합하여 분석이 가능하다.

패널자료는 미시적이고 동태적인 분석이 가능하도록 동일한 표본에 대하여 동일한 설문서 내용으로 주기적으로 조사하는 방법이다. 첫째, 사건들의 ‘인과관계 설정에 있어서 시간적 순서(time order)의 조건’을 만족시키고 둘째, 인간의 행동이나 태도, 또는 어떤 사건의 시간적인 변화의 유형이나 추세를 파악하고 설명이 가능한 분석을 할 수 있다.

다. 장단점

실증적 통계분석에 기초해서 정책 수립 및 그 효과를 측정하고자 하는 정책입안자나 연구자들은 패널조사를 선호하는 경향이 있다. 이는 횡단 조사와 시계열조사의 성격을 동시에 갖고 있는 패널조사의 장점에 기인한다(Hsiao, 2005). 구체적인 장점을 살펴보면 다음과 같다.

첫째, 모수에 대한 보다 정확한 추정이 가능하다. 패널조사는 단일시점에서 조사한 횡단자료에 비해 자유도가 높다. 그리고 시계열자료의 경우 다중공선성(multicollinearity) 문제가 발생할 가능성이 높으나 패널자료는 다중공선성 문제가 발생할 가능성이 적어진다.

보다 중요한 패널조사의 두번째 장점으로, 단일한 종단 혹은 시계열 자료에 비해 패널조사 자료를 분석하면 인간행동의 복잡성을 포착해 낼 수 있다. 먼저, 보다 복잡한 행동적 가설을 검증할 수 있다. 예를 들면, 패널자료는 사회정책 프로그램의 효과를 평가할 수 있다. 횡단자료를 이용해서 특정 프로그램 효과성을 검증하는 경우, 프로그램의 처방과 자료를 받는 사람과 그렇지 않는 사람이 본래 상이하다는 사실로 인해 어려움을 겪는다. 그러나 패널자료를 이용할 경우 동일한 개인에 대해 처방과 자료를 전후의 효과를 비교할 수 있다. 그 한 예로, 캘리포니아 주에서 범죄를 근절시키기 위해 시행한 “삼진아웃 법”을 들 수 있다. “삼진아웃 법”은 만약 용의자가 세 번의 범죄를 저지르면 평생 동안 감옥에 갇히도록 하는데, 이 법이 범죄를 단절시키는 효과가 있었는지를 평가하는 것은 법이 시행된 이후의 캘리포니아 주 범죄율을 비교해보면 된다. 이를 y_{it} 라고 하고, 같은 기간 동안 “삼진아웃 법”이 없는 다른 주 즉 오리건의 범죄율을 y_{jt} 라고 할 때, 이 둘을 직접적으로 비교하는 것은 틀린 결론을 도출할 위험이 있다. 왜냐하면, 서로 매우 다른 두 주에서 범죄율에 미칠 수 있는 여러 가지 다른 요인들이 있기 때문이다. 단순히 캘리포니아에서 법이 개시되기 이전의 범죄율 즉, y_{is} 와 이후의 범죄율을 비교하는 것도 틀린 결론을 도출 할 수 있다. 왜냐하면, 시행 이전과 이후에 기후와 인구통계가 같더라도 실업률과 같은 요인에 의해서도 영향을 받을 수 있기 때문이다. 그러나 만약 시간에 따라 변동하는 이러한 요인들이 캘리포니아와 오리건에서 유사한 형태로 변동한다면, 두 주 각각의 시행 이전과 이후의 범죄율 차이, 즉 $[(y_{it}-y_{is})-(y_{jt}-y_{js})]$ 는 범죄율에 미치는 다른 요인들의 효과를 제거한 “삼진아웃 법”의 효과를 분리해낼 수 있게 해준다. 이러한 ‘차이 차(difference in difference)’ 기법은 패널 자료에서만 가능하다(Hsiao, 2005). 즉 결과에 영향을 미치는 다른 요인을 통제된 상태에서 처치 효과만을 분리할 수 있게 된다. 또한 비관찰 혹은 생략된 변수의 효과를 통제할 수 있다. 패널자료는 시점 간 동태 정보와 전체에서 개별에 대한 정태 정보를 동시에 지니고 있기 때문에, 측정되지 않는 변수의 효과를 통제할 수 있다. 미관측 변인의 효과를 통제함으로써, 보다 엄밀한 자연 과학의 실험과 유사한 상황에서 사회경제적 변수의 효과를 측정할 수

8 농가소득의 동태적 변화요인

있게 된다. 이런 분석이 가능하기 때문에, 정책평가나 엄격한 인과적 관계 규명이 필요한 경우 자연과학의 실험연구와 가장 유사한 패널조사 설계를 선호하는 것이다.

패널의 역사가 오래된 외국에서는 이와 같은 패널자료가 가지는 특성 혹은 장점으로 인해, 정부정책의 주요 방향을 결정하는 데 패널자료가 중요한 참고자료로 활용되고 있다. 예를 들면, 독일의 가구패널조사(GSEP: German Socio-Economic Panel Study)나 영국의 가구패널조사(BHPS: British Household Panel Study)는 소득, 노동시장, 가족 등 다양한 현상에 걸쳐 그 동태적 변화를 분석하여 정책형성에 기여하고 있다(이삼식 외, 2006: 40). 패널조사 자료를 분석한 연구 또한 기하급수적으로 증가하고 있다(Hsiao, 2005). 1986년에 사회과학 인용색인(Social Science Citation index)에서 “패널 혹은 종단자료” 색인어로 29개 연구목록이 있었으나, 2004년에는 687개의 논문이 검색되었다. 이와 같은 패널자료를 이용한 연구의 기하급수적 증가는 첫째 자료의 이용가능성이 증가하고, 둘째 시계열 자료나 횡단자료에 비해 인간행동의 복잡성을 포착하는 데 패널자료가 유용하며, 셋째로 방법론적인 도전의 결과로 볼 수 있다. 자료의 이용가능성과 관련해서, 패널조사 방식이 비용이 많이 소요됨에도 선진국뿐만 아니라 개발도상국에서도 광범위하게 활용되기 때문이다.

이와 같은 복잡한 인간행동의 인과관계 규명과 관련한 패널자료의 장점에 대해 국내 여러 연구에서도 언급하고 있다. 패널자료는 경제현상 분석, 정책의 효과성, 노동시장 분석 등을 가능하게 하며, 또한 특정 기간의 효과(period specific effect), 특정 연령효과(age specific effect) 그리고 동시태생의 효과(cohort effect) 파악을 가능하게 해 준다고 지적한 바 있다(금재호, 1998). 신동균(1998), 이상호(2006) 또한 패널자료는 미관측 이질성을 통제함으로써 시간의 경과나 환경의 변화에 따른 동태적 변화 및 상태 간의 변이과정을 추적할 수 있을 뿐 아니라, 관찰 불가능한 변수들을 효과적으로 통제함으로써 엄격한 조건에서 특정 변수의 효과를 측정할 수 있다고 지적하였다.

패널조사가 지닌 이런 장점에도 불구하고 많은 패널조사가 진행되지 않는 것은 무엇보다도 현실적인 제약 때문이다. 우선, 패널조사는 동일대상을 반복해서 조사해야 하는 까닭에 패널유지를 위해 많은 비용이

소요된다. 횡단조사 설계에서는 표본을 선정할 다음에 한번의 조사로 그치기 때문에 별도의 추적비용이나 패널 관리비용이 소요되지 않는다. 또한 시간에 있어서도 패널조사 자료를 분석하기 위해서는 자료축적이 필요하다. 즉, 관심 있는 현상에 대한 의미 있는 분석을 위해서는 많은 시간을 기다려야 하는 문제가 있다. 이와 같이, 패널조사의 계획과 관리가 어려운 이유는 장기적으로 설정된 목표나 수요에 비해 단기간에 동원 가능한 자원들 사이에 항상 차이가 존재하기 때문이다. 따라서, 조사 초기부터 장기적인 목표와 자원을 충분히 고려한 상태에서 진행하지 않으면 패널조사는 지속되기 어렵다(이지연, 2005). 장기적인 관점에서 연구설계와 자료수집이 이루어지는 만큼, 조사설계는 한 번 결정되면 변경하기 어렵다. 따라서 표본추출, 설문구성 등의 사전준비를 철저히 해야 한다. 이를 위해 많은 노력과 비용이 투입되므로, 예비조사를 실시하는 것이 일반적이다. 이러한 장기 반복조사의 특성으로 인해 연구 예산 및 연구 인력의 안정적 운영이 요구된다.

마지막으로, 다른 조사설계에서는 그렇지 않으나 유독 패널조사에서 부각되는 문제점의 하나는 패널 탈락(panel attrition) 문제이다. 차수(wave)가 반복될수록 1차년도 원표본에 대해 계속 조사를 하기가 어렵다. 회상에 의존하는 횡단 조사에 비해 기억오차는 감소하지만, 패널탈락에 의한 문제가 야기된다. 특히, 표본 탈락이 비체계적으로 발생한 경우 남아 있는 패널의 대표성을 유지하기 어려운 점이 있다(신동균, 1998; 이지연, 2005).

2. 패널통계의 유형

패널통계는 조사단위, 조사의 범위, 운영 패널의 수, 표본교체 여부, 패널유지 기간 등에 따라 다양하게 그 유형을 구분해 볼 수 있다.

첫째, 대표적인 패널조사 구분방법은 조사단위에 따른 구분으로 개인패널, 가구패널, 사업체 패널 등으로 구분할 수 있다. 하지만 가구패널의 경우, 조사대상 가구원 개인에 대해서도 패널화하여 분석할 수 있기 때문에 개인 패널의 특성도 동시에 지니고 있다. 둘째, 조사의 범위에 따라 패널조사의 유형은 크게 일반형과 특수형으로 구분할 수 있다.

‘일반형’ 패널조사는 전국민을 모집단으로 하여 일반적인 사항을 전반적으로 설문하는 조사로서 우리나라의 대우패널, 미국의 PSID(Panel Study of Income Dynamics), 독일의 GSEP, 영국의 BHPS 등을 들 수 있다(국민금융연구원, 1997). ‘특수형’ 패널조사는 특정 연령대나 특정 주제를 다루는 패널로 일반패널에 비해 세밀한 연구를 수행하는 패널이다. 또한 표본프레임에 따라서는 전국단위 센서스, 학생집단, 특정지역 거주자, 기초생활 수급자 등에 따라 구분할 수 있고, 표본을 선택하는 방법에 따라 모집단 특성을 반영한 대표 패널(representative panel)과 출생시기와 같은 세대가 표본추출의 기준이 되는 코호트 패널(cohort panel)이 있다. 셋째로, 일정기간이 지난 후에 표본교체가 이루어지면 순환패널(rotating panel), 최초의 표본추출 후에 다른 표본추출이 없으면 고정패널로 구분할 수 있다. 넷째로, 동일한 패널조사의 이름으로 운영되는 패널의 수에 따라 단일패널과 복수패널로 구분해 볼 수 있을 것이다. 코호트 패널의 경우에 복수의 패널을 순차적으로 운영하는 경우가 있다. 마지막으로 패널 유지기간의 한정 여부에 따라, 기간 한정 없이 지속적으로 운영하는 경우와 종료시점을 정해 놓고 운영하는 경우가 있다.

3. 패널통계의 현황

가. 패널조사의 역사

패널조사는 1968년 미국에서 가구패널조사(PSID, Panel Survey Income Dynamics)를 실시하여 패널자료로 구축하기 시작하였고, 1984년 독일에서 GSEP, 1991년 영국의 BHPS 등 현재 주요 선진국에서는 다양한 계층의 패널조사를 구축하여 운영 중에 있다. 우리나라에 패널조사가 최초로 도입된 것은 1993년 대우경제연구소에서 시작한 일명 ‘대우패널’, 즉 한국가구패널조사(Korea Household Panel Study: KHPS)를 통해서였다. 민간 영역에서 먼저 도입된 KHPS가 IMF 시기를 맞아 중단되면서⁶⁾, 이를 이어 한국노동연구원에서 1998년 한국노동패널을 운영하면서 최초로

6) 대우패널(KHPS)은 1998년 7차 웨이브(wave)를 끝으로 중단되었다. 대우패널의 운영 경험에 기초해서 구축된 패널이 한국노동패널이다.

공적 영역에서 패널조사가 도입되었다(남재량 외, 2007). 대우패널의 경험을 이어받은 한국노동패널이 성공적인 결과를 보이자 2000년 이후 정부출연연구소, 대학, 공단을 중심으로 패널조사가 급속히 확산되었다. 2007년 현재 정부예산을 사용하여 중앙 정부부처, 국책연구기관 및 공단 등에서 운영중이거나 계획확정된 패널조사만 20여개에 달한다. 지금도 여러 정부 부처나 정부출연연구소 등에서 다양한 사회적 정책에 걸쳐 신규 패널사업을 시도하고 있어, 앞으로도 패널조사는 계속 증가할 것으로 예상된다. 또한 최근에는 민간영역의 각종 선거여론조사에서도 유권자 패널을 구축하여, 유권자의 지지정당 및 후보 지지도의 변화를 추적하는 조사설계가 이루어지기도 하고(김장수, 2007), 특히 인터넷 온라인 조사에서도 응답자 풀(pool) 구축을 통한 특정 집단 패널조사가 점차 활성화되고 있는 추세이다.

나. 국내 패널조사

1) 패널조사 현황

조사시점인 2007년 말 현재 중앙 정부부처, 정부출연연구기관 및 공단 등 공적 기관에서 정부 예산을 사용하여 현재 운영 중이거나 운영계획이 확정된 패널은 총 10개 기관의 17개로 나타났다<표 2-1>. 전체적으로 패널조사가 운영되고 있는 분야는 교육, 노동, 복지, 가족, 청소년과 관련된 정부업무로, 이는 외국에서도 패널조사가 활발히 이루어지고 있는 분야이다.

정부업무 관련 부처별로 패널조사를 살펴보면 노동부가 6개(청년패널, 대졸자 직업이동 경로, 고령화연구패널, 장애인고용패널, 한국노동패널, 사업체패널)로 가장 많았고, 다음으로 보건복지가족부 4개(인구패널, 한국복지패널, 한국의료패널, 국민노후보장패널), 교육과학기술부 2개(한국교육종단연구, 한국교육고용패널), 여성부 2개(여성인력패널, 한국아동패널), 문화체육관광부 1개(한국청소년패널)이고, 이외에 부처 공동 관련 패널로는 여성가족패널(여성부, 노동부, 보건복지가족부), 인적자본 기업패널(노동부, 교육과학기술부)이 있다.

12 농가소득의 동태적 변화요인

그리고 17개 패널을 패널운영 기관별로 보면 13개는 경제·인문사회 연구회 소속 정부출연연구소에서 운영하고 있고, 나머지 4개 패널은 한국고용정보원 2개, 국민연금관리공단과 한국장애인고용촉진공단에서 각각 1개의 패널을 운영하고 있는 것으로 나타났다. 정부출연연구소에서 직접 혹은 위탁형태로 운영하고 있는 13개 패널 현황을 보면 한국노동연구원(노동패널, 사업체 패널, 노령화연구패널)과 한국보건사회연구원(복지패널, 의료패널, 인구패널)이 각 3개, 한국직업능력개발원(교육고용패널, 인적자본 기업패널)과 한국여성정책연구원(여성가족패널, 여성인력패널)이 각 2개, 한국청소년정책연구원(청소년패널), 한국교육개발원(교육중단연구), 육아정책개발센터(아동패널)에서 각각 1개의 패널을 운영하고 있다. 이 중에서 중앙 정부부처에서 직접 운영하는 패널은 여성부의 '여성인력패널'과 보건복지가족부의 '인구패널'로 각각 한국여성정책연구원과 한국보건사회연구원에 위탁하여 조사를 진행하고 있는 것으로 나타났다. 즉, 경제·인문사회연구회 소속 정부출연연구소에서 직접 운영하고 있는 패널은 이 2개의 위탁 패널을 제외한 11개이다⁷⁾.

한편, 전체 17개 패널 중에서 통계청 승인통계가 13개, 비승인 통계가 4개인데, 비승인 통계 중에서 '장애인고용패널'과 '인구패널'은 각각 2008년과 2009년에 승인을 받을 계획이지만, '한국청소년패널'과 '여성인력패널'은 승인 신청계획이 아직 불분명한 것으로 조사되었다. 통계청의 승인을 받은 연도를 보면 노동패널이 1998년, 나머지는 대부분 2004년 이후 집중적으로 승인을 받은 것으로 나타났다. 통계청 승인년도 집중 현상도 패널조사가 짧은 기간에 집중적으로 구축된 상황을 반영하고 있다.

7) 이와 같은 예산조달 및 사업 운영주체의 문제로 인해 경제·인문사회연구회에서 통계 가능한 사업도 11개 패널로 한정되어, 패널이라는 관점에서 전체 패널사업을 조정할 수 있는 기관이 부재한 무정부주의 상황이다.

〈표 2-1〉 국내 패널조사

번호	패널 조사명	운영기관	관련 정부부처	국가통계 승인 (승인년도,번호)
1	한국청소년패널조사 (KYPS: Korea Youth Panel Survey)	한국청소년 정책연구원	문화체육 관광부	비승인 (계획 없음)
2	한국교육종단연구2005 (KELS 2005 : Korea Education Longitudinal Study)	한국교육개발원	교육인적 자원부	승인 (2006, 33408)
3	한국교육고용패널조사 (KEEP: Korea Education & Employment Panel)	한국직업능력 개발원	교육인적 자원부	승인 (2004, 38902)
4	청년패널조사 (YP : Youth Panel)	한국고용정보원	노동부	2007, 32705
5	대졸자 직업이동 경로조사 (GOMS: Graduates Occupational Mobility Survey)	한국고용정보원	노동부	승인 (2006, 32704)
6	여성인력패널조사	한국여성정책연구원 (여성부 수탁)	여성부	비승인 (잘 모름)
7	고령화연구패널조사 (KLoSA: Korea Longitudinal Study of Aging)	한국노동연구원	노동부	승인 (2006, 33602)
8	장애인고용패널조사	한국장애인고용 촉진공단	노동부	승인 (2007, 38303)
9	한국아동패널 (Panel Study of Korean Children)	육아정책개발센터	여성부	비승인 (08년 예정)
10	인구패널	한국보건사회연구원 (보건복지가족부 수탁)	보건복지가족부	비승인 (09년 예정)
11	한국노동패널조사(KLIPS)	한국노동연구원	노동부	승인 (1998, 33601)
12	여성가족패널조사 (KLoWF : Korea Longitudinal Survey of Woman & Family)	한국여성정책연구원	여성부/ 노동부, 보건복지가족부	승인 (2007, 33801)
13	한국복지패널조사연구 (KoWePS: Korea Welfare Panel Study)	한국보건사회연구원	보건복지가족부	승인 (2006, 33109)
14	한국의료패널 (KoHPS : Korea Healthcare Panel Study)	한국보건사회연구원	보건복지가족부	승인 (2007, 33110)
15	국민노후보장패널조사 (KRIS: Koeran Retirement & Income Study)	국민연금관리공단 (국민연금연구원)	보건복지가족부	승인 (2005, 32201)
16	사업체패널조사 (WPS : Workplace Panel Survey)	한국노동연구원	노동부	승인 (2006, 33603)
17	인적자본 기업패널조사 (HCCP : Human Capital Corporate Panel)	한국직업능력개발원	노동부, 교육과학기술부	승인 (2005, 38903)

2) 패널조사의 유형

조사대상 패널에 대해 몇 가지 기준을 적용하여 그 유형을 나눠보면, 이미 다양한 종류의 패널이 공적 영역에서 운영되고 있음을 알 수 있다 <표 2-2>. 첫째, 조사단위별로 보면, 단위가 개인인 패널이 8개, 가구단위 조사 7개, 사업체 단위 조사가 2개로 다양한 조사단위의 패널이 운영 중이다. 개인단위 조사의 경우 생애주기가 청소년 코호트인 경우가 청소년패널, 교육종단연구, 교육고용패널, 청년패널, 대졸자 직업이동 패널 등 5개로 대부분을 차지하고 있다. 가족으로부터 일정 정도 분리된 청소년에 대한 조사는 가구조사보다 청소년 개인단위 조사 설계를 선호하는 것으로 보인다. 하지만, 개인단위 조사의 경우에도 1차 조사에서는 가족이나 부모에 대한 기본사항을 파악하는 경우가 많아, 순수하게 청소년 개인만을 조사하는 것은 아니다. 청소년패널의 경우와 같이 최초 조사 시에는 청소년의 학부모에 대해 조사를 실시하는 경우도 있고, 학교효과를 파악하기 위해 단위학교도 패널방식으로 추적조사를 하고 있는 교육종단연구의 경우도 있다.

가구단위 조사의 경우 개인보다는 가족상황에 의해 영향을 많이 받는 아동, 인구, 노동, 복지 등 다양하고 복합적인 주제를 다루는 경우가 많다. 가구단위 조사의 경우에도 조사대상 가구원에 대한 개인패널 구축이 가능하기 때문에, 개인단위 조사와 엄격히 구분되지는 않는다. 이런 경향으로 인해 많은 가구단위 패널조사 운영자들은 ‘개인 및 가구단위’ 조사로 응답한 경우가 많았다. 개인에 분석의 초점을 맞추는 개인단위 조사에 비해 가구단위 조사는 가구차원의 분석이 가능하고, 추적 조사단위가 가구라는 특성을 지닌다. 한편, 조사단위가 사업체인 패널은 2개로 사업체패널과 인적자본기업패널이 있다.

둘째, 단일 사업 명칭 하에 실제 운영 패널이 단일패널인 경우가 17개 중 15개로 대부분이고, 2개의 패널이 운영되는 복수패널은 청소년패널과 청년패널이다. 교육고용패널의 경우, 같은 시기에 중학교 3학년과 일반계·전문계 3학년을 별도의 표본으로 구성하고 있어 실질적으로는 복수패널로 운영되고 있다. 교육종단연구, 교육고용패널, 대졸자 직업이동 패널 등은 향후 복수의 패널로 운영될 가능성이 높을 것으로 판단된다.

청소년 코호트 패널의 경우, 생애주기가 일정기간 경과하면 청소년패널과 같이 추가 패널을 구축하는 경우가 많기 때문이다.

셋째, 패널운영 기간이 한정되어 있는 기간한정패널과 지속운영패널을 구분해서 보면, 청소년패널(중2 패널 2008년 종료), 대졸자직업이동(8년 차), 교육중단패널(2023년) 및 노후보장패널(2015년)만 패널운영 만료기간을 명시하고 있을 뿐 나머지 패널은 최소 운영기간만 명시하거나 무기한, 미정인 경우가 많아 지속운영 여부에 대해 명확한 계획이 없는 경우가 대부분이다. 이는 가능한 장기적으로 패널을 운영할 계획을 보여 주는 것으로, 생애 특정 코호트 조사의 경우 복수의 패널을 운영하거나 혹은 표본개편 등을 통해 조사를 지속할 것으로 예상된다. 예를 들면, 교육고용패널은 학교에서 직업으로의 전환에 주된 목적을 두고 있음에도 최소 13년 이상 조사를 지속할 계획이고, 청년패널도 10년 이상으로 명확한 만료시점을 제시하지 않고 있다. 고령화연구패널, 노동패널, 사업체패널은 지속조사 혹은 무기한 조사라고 응답해, 종료계획이 없는 것으로 파악된다. 이와 같은 불명료한 조사만료 시점은 향후 패널사업 예산 증가의 중요한 원인 중 하나로 작용할 것이다.

넷째, 조사주기를 보면 13개 패널이 1년 주기로 조사를 실시하고 있고, 나머지 4개 패널(청년패널, 노후보장패널, 사업체패널, 인적자본기업패널)만이 2년 주기로 조사를 실시하고 있다. 사업체 단위 2개 조사는 모두 2년 주기이고, 청년패널의 경우 2007년부터 1년에서 2년으로 전환하였으며, 여성가족패널의 경우 조사 안정화 이후에는 2년 주기로 전환할 계획인 것으로 응답하였다. 한편 1년 주기인 고령화패널은 기본조사 와 특별조사를 격년제로 시행하고 있는 특징을 보이고 있다. 교육중단 연구는 단계별로 1단계인 2010년까지 1년 주기 운영 후, 2단계(2011-19년)는 2년 주기, 3단계(2020-23년)에는 다시 1년 주기로 운영할 계획이다.

16 농가소득의 동태적 변화요인

<표 2-2> 패널 설계유형

번호	조사명	조사단위	복수패널 여부	조사주기	완료예정
1	청소년패널	개인	중2 패널	1년	2008
			초4 패널	1년	
2	교육종단연구	개인	단일	1단계(01-10년): 1년, 2단계(11-19년): 2년, 3단계(20-23년): 1년	2023
3	교육고용패널	개인	단일	1년	최소 13년
4	청년패널	개인	YP2001	2006년 1년주기, 이후 2년주기	10년이상
			YP2007	1년주기	10년이상
5	대졸자직업이동	개인	단일	1년	8년차까지
6	여성인력패널	개인	단일	1년	미정
7	고령화연구패널	개인	단일	1년 (기본, 특별 각 1년)	없음
8	장애인고용패널	개인	단일	1년	미정
9	아동패널	가구	단일	1년	2019
10	인구패널	가구	단일	1년	미정
11	노동패널	가구	단일	1년	지속 조사
12	여성가족패널	가구	단일	1년(안정 후 2년주기)	
13	복지패널	가구	단일	1년	미정
14	의료패널	가구	단일	1년	미정
15	노후보장패널	가구	단일	2년	2015
16	사업체패널	사업체	단일	2년	무기한
17	인적자본기업패널	사업체	단일	2년	최소 10년

3) 조사 대상과 목적 및 내용

조사목적 및 내용을 파악하기 위해 조사목적, 조사내용, 패널조사 방식으로 설계한 이유 및 연구와 관련된 정책이슈 등에 대해 질문하였다. 여기서는 17개 패널이 다루고 있는 조사목적, 조사내용과 함께 조사대상도 동시에 검토하고, 이를 기초로 패널조사 간 유사중복 문제도 함께 다루고자 한다<표 2-4>.

전체 패널조사의 조사대상을 생애주기 분류<표 2-3>로 보면 출산과 관련된 인구패널, 신생아를 대상으로 하는 아동패널, 청소년을 대상으로

하는 청소년패널, 교육중단연구, 청년패널, 교육고용패널, 학교 졸업 후에는 대졸자 직업이동패널, 여성인력패널, 노동패널이 있고, 중고령자에 대해서는 고령화 연구패널, 노후보장 패널이 있다. 출생에서 노후까지 모든 연령층을 포괄하여 전 생애에 걸쳐 패널조사가 진행되고 있다.

〈표 2-3〉 생애주기별 분류

유아기	청 소 년 기			대 학	청 년	중 년	노 년	
	초 등	중 등	고 등					
육아 패널	청소년 패널	청소년패널, 교육패널, 교육고용 패널	교육패 널, 교육고 용패널	교육패널, 교육고용 패널	청년패널, 대졸자패널		고령화 패널	
							국민노 후보장 패널	
							노동패널	
							복지패널	
		여성가족패널						

특히, 청년패널, 교육고용패널, 교육중단연구, 청소년패널, 대졸자 직업이동 패널 등이 동시에 운영되고 있어 청소년 연령대를 다루는 여러 개의 패널이 존재한다는 특징을 보여준다. 이와 같은 청소년 코호트 패널이 다양하게 존재하는 이유는, 물론 청소년기가 다양한 변화양상을 보인다는 특징도 있겠지만, 본격적으로 패널조사가 도입된 2000년 전후에 IMF 경제위기로 인한 청년 실업 문제가 심각한 사회적 이슈로 등장했다는 사실과 관련이 있는 것으로 추측된다. 이 외에도, 우리 사회의 당면한 사회문제를 중심으로 패널이 구축되는 경향을 확인할 수 있는데, 저출산 문제와 관련해서 인구패널이, 급속한 고령화 문제를 다루는 고령화 연구패널과 노후보장 패널이 운영되고 있다. 경제활동과 관련해서는 노동패널, 장애인고용패널, 복지패널, 여성가족패널 등이 운영되고 있다.

한편, 패널조사에서 다루고 있는 조사내용을 보면 교육, 학교에서 직업으로의 전환, 노동, 여성, 아동, 인구, 복지, 의료, 사업체를 대상으로 노사관계, 인적자본 등 다양한 분야에 걸쳐 패널조사가 진행되고 있다.

조사단위에 따라서 개인단위 조사는 학교·가정·사회생활, 사교육, 진로·직업교육훈련, 결혼·출산, 경제활동, 사회보장 등을 주로 다루고 있다. 가구단위 조사의 경우, 가구변동, 가족관계, 주거환경, 보육실태, 가족경제 등을 주로 다룬다. 사업체 단위 조사는 사업체의 특성, 노사관계, 인력개발 등에 관심을 가지고 있다.

개별 패널에 대해 조사 대상, 목적, 패널설계 이유, 관련 정책 이슈 및 조사내용을 중심으로 살펴보고자 한다. 먼저 고령화연구패널과 인구패널을 제외한, 나머지 6개의 개인단위 조사는 모두 청소년과 청년층을 대상으로 학교 및 직업세계를 다루고 있는 특징을 보인다. 청소년패널은 조사시점에서 중학교 2학년과 초등 4학년 학생집단을 대상으로, 청소년의 진로설정 및 준비, 여가, 일탈, 생활영역, 자아관 등을 조사하여, 청소년에 대한 종단적 데이터 구축을 주된 연구목적으로 삼고 있다. 관계된 정책이슈로는 청소년의 진로의식 및 진로과정(고졸청소년의 노동시장 진입, 학업성취 및 대학 진학), 일하는 청소년의 변화와 사교육 참여, 청소년의 일탈행위 실태 및 원인, 여가제약 및 여가참여 양상, 심리상태 및 부모·교사·친구관계 등이다. 청소년 연령대의 성장과정에 따른 변화양상과 인과관계 파악을 위해 종단적 연구설계로 운영 중이다.

교육종단연구의 경우, 중학교 1학년을 대상으로 교육성취(학생성취도 검사 및 교사 평가), 가정생활, 학교생활, 학습활동, 여가, 자아개념 등을 조사하고, 학생들이 경험하는 학교교육을 매년 조사하여 그러한 경험이 학생들의 인지적·정의적 발달에 어떤 영향을 미치는가를 파악하고자 한다. 특히 이 과정에서 가정적 배경에 따라 학교교육의 경험과 결과에 어떤 차이가 있으며, 학교의 종류와 유형이 어떤 영향을 미치는가를 파악하는 것이 주된 목적이다(한국교육개발원, 2005). 관련된 정책 이슈는 학교 폭력 및 학생 일탈, 사교육 및 과외 경감 대책, 학교정책의 효과, 평가 및 개선방안, 고교평준화 제도의 논란(하향평준화), 전문계 교육의 부실화와 발전방향, 교육격차 해소를 위한 방안 등이다. 학교 교육 및 정책 효과 평가를 위해서 동일 학생을 반복 측정하는 조사방법을 활용하고 있다.

교육고용패널은 청년층의 교육관련 정보구축을 통해 진학, 진로, 노동시장 진입 관련 연구수행 및 학교교육과 직업훈련의 성과가 경제적

성과에 기여하는 정도를 파악하는 데 조사의 초점이 두어져 있다. 이를 위해 중학교 3학년과 일반계 및 전문계 고등학교 3학년을 대상으로 학교 및 가정생활, 여가, 사교육, 근로경험, 진로계획 및 진로지도 등을 조사하고 있다. 그리고 학생들에게 영향을 미치는 가정 및 학교환경을 조사하기 위해 학생의 보호자와 담임교사, 학교행정가를 동시에 조사한다. 관계된 정책적 관심사는 교육-노동시장 연관성 파악, 노동시장 정책의 효과성 검증, 기업 인적자원 개발 및 활용 기여, 자기개발 및 평생직업 교육의 방향 제공 등이다. 청년층 인적자원의 현재 수준 및 축적과 활용 과정을 파악하여 인적자원개발 정책수립에 기여하기 위해 종단적인 패널조사로 구축하였다.

청년패널은 청소년의 학교생활, 사회·경제활동, 가정배경 등의 자료를 수집하여 청년실업해소를 위한 노동정책 수립 및 연구를 위해, 한국고용정보원에서 구축한 패널이다. 공적인 패널조사 중 국내 최초의 개인패널로, 15~29세 청소년을 대상으로 2001년에 조사를 시작하였고, 조사의 초점이 학교에서 직업으로의 이행과정 및 노동시장 경로(career path)에 두어져 있다. 청년실업, 인력수급의 양적·질적 불일치, 사교육·교육훈련·자격의 효과 등이 이 조사의 정책적 관심사항이다.

대졸자 직업이동조사는 2년제와 4년제 대학 졸업생을 대상으로 학교 교육, 직장경험, 직업훈련 등을 파악하여 대학졸업자의 경력개발 및 직업 이동 경로를 추적하고 있다. 조사대상은 대학졸업자(2004년 8월 및 2005년 2월)의 5%인 25,000명으로 교육과정, 현재 일자리, 졸업 후 일자리, 일자리 탐색과정, 재학 중 취업, 훈련과정 및 자격증 등 노동시장 진입 및 정착에 영향을 줄 수 있는 요인들에 대해 주로 질문하고 있다. 관련된 정책적인 쟁점은 고학력 청년실업 문제, 직업불일치 문제, 과잉교육, 직업교육 및 능력개발, 여성의 경력단절 등이다. 패널조사로 설계한 이유는 특정연도 졸업생의 노동시장 진입과 정책, 경력개발 등의 정보를 얻기 위해 동일인의 장기적인 변화 파악이 필요했기 때문이다.

앞서 살펴본 청소년 코호트 패널들과는 달리, 여성인력패널과 장애인 고용패널은 여성관리자와 장애인이라는 특정 집단을 대상으로 한다는 점에서 차이가 있지만 이들의 노동시장 경력형성에 관심이 있다는 점에서는 유사하다. 여성인력패널은 기업 내 여성관리자의 경력형성 과정을 추적

하고, 여성관리자의 환경요인과 이들이 여성관리자에게 미치는 영향을 파악하기 위해, 대리 2년차 이상 여성관리자를 대상으로 직장 근무사항, 교육 및 직업훈련, 경력개발 및 이동사항 등을 조사하고 있다. 추가로 기업인사담당자에 대해 기업현황, 인사관리제도, 인력운용 현황 등도 파악하고 있다. 이 조사와 관련된 정책적 쟁점은 국가성평등지수 개선 기여, 여성의 관리직 진출 장애요인 제거이고, 여성 개인의 경험을 추적하여 상이한 경력형성 과정을 파악하기 위해 종단적 연구설계로 진행하고 있다.

장애인고용패널은 법정장애유형을 지니고 있는 재가장애인을 대상으로, 장애인의 취업 및 직업유지에 영향을 미치는 개인적, 환경적 요인을 규명하고자 한다. 조사내용을 보면 경제활동상태, 장애정보, 직업능력, 고용서비스, 직업능력개발, 일상생활 및 삶의 질에 초점을 맞추고 있다. 장애인의 의학적 손상과 직업적 능력 및 고용성과와의 관계, 생애주기별 경제활동상태와 삶의 질, 생산적 복지 구현과 체계적인 직업재활시스템의 구축과 평가, 고용확대방안 모색이 주된 정책적 관심사이다. 즉, 장애인의 장기간에 걸친 경제활동상태 및 직업상의 변화와 이동과정을 분석하기 위해 종단적 연구설계를 활용하고 있다.

청소년 및 청년 코호트와 여성인력 및 장애인 대상 개인 패널을 살펴본 결과, 두 가지 공통적인 사항을 확인할 수 있다. 첫째, 이들 조사의 공통적인 관심이 학교 및 직업세계에 주어지고 있다는 점이다. 조사에 따라 교육효과, 인지적·정의적 발달, 문제행동, 사교육, 장애인 환경, 여성관리자의 장애 요인 등 세부적인 내용과 깊이에서는 차이를 보이지만, 공통적 관심사는 청소년과 여성 및 장애인의 중요한 발달과제인 학교에서 직업세계로의 전환 및 적응 문제이다. 이와 같은 공통의 관심사로 인해, 특히 청소년 코호트 패널은 유사중복 통계의 가능성이 지적되었다. 이런 지적으로 교육고용패널, 교육종단연구, 청소년패널의 경우 조사대상과 조사내용에서의 중복가능성으로 사전협의를 통해 연구내용을 조정할 바 있다(경제·인문사회연구회, 2007). 청소년 코호트 패널의 유사중복 문제는 별도로 검토될 것이다.

둘째, 패널조사 방식으로 설계한 이유를 보면 동태적 변화과정 추적, 청소년 및 특정 집단과 관련된 문제나 정책효과 검증 및 인과적 연구 등을 공통적으로 지적하고 있다는 점이다. 횡단조사나 횡단면 반복조사

에 비해 패널조사가 동일 대상에 대한 동태적 변화를 추적하는 조사설계이고, 특정 정책이나 가족과 학교, 직업 등의 인과적 효과를 가장 체계적으로 검증할 수 있기 때문에 패널조사를 선호하고 있다는 것이다. 이는 개인뿐만 아니라 가구 및 사업체 패널에서도 동일하게 확인된다. 개별 패널마다 우리 사회의 중요한 정책적 관심사항을 적시하고 있어, 실제 연구결과 및 정책적 활용에 대해서는 별도로 살펴볼 것이다.

이어서 가구단위 조사를 보면 노동패널을 시작으로 여성가족패널, 인구패널, 복지패널, 의료패널, 아동패널, 노후보장패널 등 다양한 분야에 걸쳐 패널조사가 운영되고 있다. 출생아동과 그 부모를 취학 전까지 추적하는 아동패널은 기존 패널과 변별되는 연구대상과 주제를 다루고 있다. 아동의 발달과정과 부모의 자녀양육 분야는 맞벌이 가정의 양육 문제, 유아교육과 보육의 통합이라는 정책적 과제와 연결되어 있다. 한편, 여성가족패널은 여성의 경제활동 참여촉진 및 경제적 자립, 여성빈곤화 방지 등과 같은 가족정책수립을 위한 기초자료 제공을 목적으로 구축된 패널이다. 이를 위해 19~64세 여성가구를 대상으로 생애주기별 경제활동상태와 가족구조, 가족과 관련된 가치, 가족관계, 가족과정 및 가족구조의 변화 추세를 추적한다. 그리고 인구패널은 저출산 문제에 대한 대책을 마련하기 위해 구축된 패널로, 전국 표본가구의 미혼자, 기혼자, 청소년을 대상으로 출산 관련 태도, 정부정책에 대한 수용도, 출산에 영향을 미치는 요인을 파악하고자 한다.

노동패널은 대표적인 가구패널로 도시거주 표본 가구의 15세 이상 가구원을 대상으로 개인의 경제활동, 노동시장 이동 및 성과를 장기적으로 추적하는 조사이다. 가구의 소득과 소비, 자산과 부채, 사교육, 주거 등 다양한 영역을 다루고 있고, 개인에 대해서는 경제활동상태, 일자리 특성 및 만족도, 근로시간 및 임금, 교육훈련 등을 주로 조사한다. 이를 통해 실업의 성격 규명, 노동력 상태변화 및 소득이동 등을 파악하고자 종단적 연구설계로 진행하고 있다. 한편 복지패널은 표본가구를 일반가구와 저소득가구로 구분하여 표집하고, 표본가구의 15세 이상 가구원을 대상으로 일반사항, 가계수지, 경제활동, 사회복지제도와 관련된 항목을 주로 질문한다. 빈곤과 사회양극화, 비정규직 확대와 청년실업문제, 가계부채 등의 문제가 복지패널과 관련된 정책적 이슈이다.

〈표 2-4〉 조사목적 및 정책 이슈

조사명	조사대상	조사목적	관련 정책 이슈
한국청소년 패널[KYPS] [한국청소년 정책연구원]	· 표본틀: 2003 교육통계연보 - 중 2 : 3,000명 (학생, 학부모) - 초 4 : 2,500명 (학생, 학부모)	- 청소년들의 직업 선택, 향 후 진로 설정, 진로 준비, 여가, 일탈 행위 등에 대한 종단적 데이터 구축	- 진로의식 및 진로과정 (고졸청소년의 노동시장 진입, 학업성취 및 대학 진학) - 일하는 청소년의 변화와 사 교육 참여 - 청소년 일탈행위 실태 및 원인 - 여가제약 및 여가참여 양상 - 심리상태 및 부모/교사/ 친구관계
한국교육 종단연구 [KELS] [한국교육개 발원]	· 표본틀: 2005 교육통계연보 · 중학교(150개교) 1학년 6,908명 *학부모, 담당교 사, 학교장, 학 교통계 등 조사 *대체표본 추가 표집 · 중1학생을 17년 간 추적조사	- 학생 개개인이 여러 수준 의 교육단계를 거쳐 직업 세계로 진입해 가는 과정 을 추적 조사하여 주요 교 육문제와 진로문제에 대한 자료 수집, 분석 - 중등교육 관련 종단적 자 료의 수집, 정책 수립의 기 초 자료 수집, 인적자원의 육성 및 관리를 위한 자료 수집 - 학교 교육의 효과성, 사회 이동과정에서 학교교육의 영향력에 초점	- 학교 폭력 및 학생 일탈 - 사교육 및 과외 경감 대책 - 학교정책의 효과 및 평가, 개 선방안 - 고교평준화 제도의 논란(하향 평준화) - 전문계 교육의 부실화와 발 전방향 - 교육격차 해소를 위한 방안
한국교육 고용패널 [KEEP] [한국직업능 력개발원]	· 표본틀: 2003 교육통계연보 · 중 3: 2,000명 · 일반고 3, 전문 계고 3: 각 2,000명 *보호자, 담임교 사, 학교행정가 동시 조사	- 중등교육에서 고등교육으 로 진학하는 과정과 중등 및 고등교육에서 직업세계 로의 이행과정에 대한 기 초자료 제공 - 학교교육의 투입, 과정, 산 출에 대한 요인 조사, 학교 교육 요인이 노동시장으로 진입과 이동에 미치는 요 인 분석	- 교육-노동시장 연관성 파악 - 노동시장 정책의 효과성검증 - 기업 인적자원 개발 및 활용 기여 - 외국과의 비교 및 정보교류의 장 확대 - 자기개발 및 평생직업교육의 방향 제공
청년패널 [Youth Panel] [한국고용정 보원]	· 표본틀: 1995 인구주택총조 사의 10% · 만 15~29세의 청년층 5,900명 (2001년 기준)	- 청년층의 학교에서 직업으로 이행과정 및 노동시장 경로 (career path)에 대한 광범위 한 자료획득	- 청년실업 - 인력수급의 양적/질적 불일치 - 사교육/교육훈련/자격의 효과

〈표 2-4〉 계속

조사명	조사대상	조 사 목 적	관련 정책 이슈
대졸자 직업 이동 경로 조사 [한국고용정 보원]	· 표본틀: 전문대, 대학졸업생 DB · 전문대/대학 졸 업생 25,000명 (2004.9/2005.2 졸업)	- 대학졸업자의 경력개발 및 직 업이동경로를 추적 조사하여 DB로 구축함으로써 교육·노 동 시장간 신뢰성 있는 인력 수급 정보 제공 및 인력수급 불일치 완화 도모를 위한 정 책수립 기초자료로 활용	- 고학력 청년실업 문제 - 직업불일치 문제 - 과잉교육, 직업교육 및 능력개발 - 여성의 경력단절
여성인력 패널 [한국여성 정책연구원]	· 기업내 여성관리 2년차 이상 관리자 · 기업인사담당자	- 기업내 여성관리자의 경력형성 과정 추적, 여성관리자의 환경 요인과 이들이 여성관리자에게 미치는 영향을 파악함	- 국가성평등지수 개선 기여 - 여성의 관리직 진출 장애요인 제거
고령화연구 패널 [한국노동 연구원]	· 표본틀: 2005 인 구주택총조사의 10% 표본조사구 261,237개 · 45세 이상 개인 약 10,000명(제주도 제외 전국 6,171 가구) *질병관리본부와 연계, 검진조사 예정	- 중고령인구의 정확한 실태조사 를 통해 향후 고령사회로 변화 해 가는 과정에서 개인의 행동 을 예측하고 이를 토대로 효과 적인 사회경제정책을 수립하고 시행하는데 활용될 기초 자료 생산	- 중고령자 경제활동참 여 및 소득과 자산 - 사회복지욕구, 건강 및 의료시설 이용 - 중고령자의 가족관계
장애인고용 패널 [한국장애 인고용촉진 공단]	· 법정장애유형을 지니고 있는 전 국의 만 15세 이 상 재가장애인	- 장애인의 경제활동 관련 전반 적인 기초통계를 생산하고, 취 업과 직업유지에 영향을 주는 개인적, 환경적 요인을 규명하여 고용정책의 수립 및 평가에 필요한 통계자료 생성	- 장애인의 의학적 손 상과 직업적 능력 및 고용 성과와의 관계 - 생애주기별 경제활동 상태와 삶의 질 - 생산적 복지 구현과 체계적인 직업재활시 스템 구축과 평가/ 고용확대 방안 모색
한국아동 패널 [육아정책 개발센터]	· 2008년 신생아 (1,800명) · 신생아 일차양육 자(주로 어머니) · 신생아 어머니 · 신생아 아버지	- 출생에서부터 취학 전까지 아 동의 성장, 발달 및 생활에 관 한 종단적 자료 수집 - 부모, 가정 및 육아지원을 포함 하는 아동의 주요 환경으로부터 제공되는 경험과, 이러한 경 험이 단기적인 성장·발달에 미치는 영향, 장기적으로 초등 학교에서의 적응 및 성취에 미 치는 영향 파악	- 맞벌이 가정의 대리 양육 서비스 - 유아교육 및 보육 통 합과 관련된 기초자 료 제공

24 농가소득의 동태적 변화요인

<표 2-4> 계속

조사명	조사대상	조사목적	관련 정책 이슈
인구패널 [한국보건사회연구원]	· 전국 12,000 가구의 미혼자, 기혼자, 청소년	- 저출산대책 마련을 위해 출산에 영향을 미치는 요인에 대한 관계 규명 등 - 출산에 대한 인식이 어떤 과정을 거쳐 형성되고 출산에 어떤 영향을 미쳐 결과로 나타나는지를 파악	- 저출산 지속에 따른 인구노령화 - 장기적 인구감소에 따른 사회경제적 문제
한국노동패널 [한국노동연구원]	· 우리나라의 도시 거주 5,000 가구와 해당 가구의 15세 이상 가구원	- 가구와 개인의 경제활동, 노동시장 이동 및 성과를 장기간 추적 조사하여 동태적으로 파악함으로써 횡단면 자료만으로 불가능한 심도 있는 정책수립을 위한 기초자료 제공	- 실업의 성격, 노동력 상태변화, 건강상태 변화 등을 파악
여성가족패널 [한국여성정책연구원]	· 전국 8,500가구만 19~64세 여성 (약 10,000명)	- 여성의 생활세계와 가족의 구조 및 변화실태를 파악하기 위한 패널조사 - 여성의 생애주기별 경제활동 지위 변화와 가족생활과의 관계를 비롯하여, 가족가치/가족관계의 변화, 가족과정과 이벤트의 변화, 가족구조의 변화를 횡단면뿐만 아니라 종단면적으로 추적	- 여성의 경제활동 참여 촉진 및 경제적 자립 증진 여성빈곤화 방지 - 가족정책수립의 기초자료
한국복지패널 [한국보건사회연구원]	· 전국 표본가구에 속한 15세 이상의 가구원 - 가구(7,072명) - 가구원(14,463명) - 아동부가(759명)	- 저소득가구와 일반가구에 대한 대표성 있는 표본을 확보하여 조사의 정확성과 역동적 분석의 유의성을 동시에 확보 - 체계적인 조사표 작성을 통해, “복지패널”로서의 정체성을 확보함으로써 사회복지 관련지표 생산과 관련 연구·분석 등에 널리 활용	- 빈곤과 사회양극화의 심화 - 비정규직 확대와 청년실업의 증가 - 가구소비의 위축과 가계부채의 증가

<표 2-4> 계속

조사명	조사대상	조 사 목 적	관련 정책 이슈
한국의료 패널 [한국보건사 회연구원]	· 전국 약 8,000가구	<ul style="list-style-type: none"> - 보건의료 이용과 비용지출의 수준 및 배분을 추정하고, 의료전달시스템 및 보험체계의 동태적 변화에 대한 데이터베이스 구축 - 의료이용 및 의료비 패널데이터 구축을 통한 실증자료를 바탕으로 보건의료정책의 수립·시행 - 조사모형은 의료이용 및 의료비 지출에 영향을 미치는 요인들을 파악할 수 있는 통계와 심층적인 의료이용 및 의료비 지출 내용을 포함하는 모형 구축 - 비급여 등 본인부담관련, 의약품 지출, 민간보험관련 통계생산 	<ul style="list-style-type: none"> - 보건의료전달체계 분석 - 의료비의 데이터베이스 구축 및 민간보험 규모 및 내용 파악 - 의료이용 행태 분석
국민노후 보장패널 [국민연금관 리공단]	<ul style="list-style-type: none"> · 표본틀: 2000 인구주택총조사 의 10% 표본조사구 24,995개 · 만 50세 이상 가구원 거주 5,133가구 	<ul style="list-style-type: none"> - 노후소득보장의 욕구와 중요도 파악, 연금 및 고령화 정책 수립을 위한 기초자료 생성 	<ul style="list-style-type: none"> - 인구고령화와 중/고령자의 노후준비 - 공적연금 및 노후소득보장정책 - 중/고령자의 경제적 복지
사업체패널 [한국노동연 구원]	<ul style="list-style-type: none"> · 농림어업 및 광업을 제외한 전산업에서 상용근로자 30인 이상인 일반사업장과 공공기관 (일반사업장 : 3,916개, 공공기관 : 359개) 	<ul style="list-style-type: none"> - 사업체의 동태적 변화와 고용구조변동에 대한 정보를 체계적으로 파악하여, 기업의 고용·평가·보상·교육·훈련 등 전반적 인적자원관리 및 노사관계 체계를 평가하고 새로운 사회경제적 환경에 적합한 인적자원관리 및 노사관계 모델 및 정책을 개발하는 데 기초적인 정보를 제공 	<ul style="list-style-type: none"> - 비정규직 고용의 실태와 현황, 전망 - 인적자원관리 정책 - 노사간 갈등과 대립, 노사관계
인적자본 기업패널 [한국직업능 력개발원]	<ul style="list-style-type: none"> · 근로자수 100인 이상이면서 일반기업 이상 - 기업 : 450개 - 사업장 : 205개 - 근로자 : 14,631명 (팀장 4,338명 및 팀원 10,293명) 	<ul style="list-style-type: none"> - 한국의 기업이 지닌 인적자원을 파악하고, 기업들이 인적자원개발에 기울이는 노력, 그리고 이 같은 기업의 인적자원 및 인적자원개발이 기업의 성과에 미치는 영향에 대한 시계열 자료를 구축함 	<ul style="list-style-type: none"> - 기업의 인적자원관리 및 개발과 근로자의 지식 및 숙련 형성

4) 표본추출

조사대상 패널의 표본추출 방법을 보면, 개인 및 가구단위 15개 조사의 모집단은 노동패널만 도시거주 가구이고 나머지는 특정 조건을 충족하는 전국단위임을 알 수 있다<표 2-5>. 표집틀(sample frame)은 개인패널의 경우 학생대상 조사는 해당연도 교육통계연보를 표집틀로 이용하는 경우가 대부분이다. 예외적으로, 개인단위 패널에서 대졸자직업이동의 경우 고등교육기관 졸업자 취업통계조사, 청년패널은 산업직업별고용조사, 여성인력패널은 한국신용평가에 포함된 100인 이상 기업을 표집틀로 사용하고 있다. 가구단위 패널의 경우 모두 인구주택총조사를 표집틀로 하고 다단계층화 확률표집 방법을 활용해서 표본을 추출하고 있는데, 아동패널만 일정 분만건수(상반기 250건)가 있는 산부인과 병원을 표집틀로 사용하고 있다. 이는 적합한 표집틀이 존재하지 않기 때문으로 추측된다. 사업체 단위 패널의 경우 사업체패널은 사업체기초통계조사와 공공기관(기획예산처 기준), 인적자본 기업패널은 한국신용평가(2005)를 표집틀로 사용하고 있다.

표집 방법을 보면 주로 확률표집법을 사용하고 있는데, 다단계확률 표집법을 대부분 활용하고 있다. 즉 여러 단계에 걸쳐 층화표집, 집락표집, 계통표집법을 사용하고 있다. 패널별 세부적인 표집방법을 검토하는 것은 별도의 심층 연구가 필요하기 때문에 여기서는 몇 가지 예외적인 경우만 살펴보고자 한다. 청년패널은 산업직업별 고용구조조사(OES)의 부가조사 형식을 취한다. 별도의 표본 추출틀 없이 OES에서 조사된 5만 가구 중 조사구별 무작위 추출을 통해 20%인 1만 가구를 선정한 다음, 15~29세 인구비율을 25%로 추정하여 8,000명을 표본으로 선정하였다. 이와 같이 OES의 부가조사 방식의 이중추출법을 활용하였기 때문에 지역별, 연령대별 패널분포가 모집단과 차이가 있을 수 있다. 대졸자직업이동의 경우에도 한국교육개발원의 졸업자취업통계조사에서 파악한 전국 전문대 이상 졸업생을 모집단으로 하여 다단계층화추출법을 사용하고 있다.

비확률 표집법을 이용하고 있는 패널로는 여성인력패널, 아동패널, 인적자본기업패널이 있다. 여성인력 패널의 경우 전국 100인 이상 기업의

<표 2-5> 표본추출

ID	조사명	복수 패널	모집단	표집틀	표집방법	조사대상	표본수
1	청소년 패널	중2	전국 중2 청 소년(2003)	해당년도 교육통계연보	층화다단계 집락표집	중2 학생, 학부모(최초)	3,449명
		초4	전국 초4 청 소년(2004)			초4 학생, 학부모 (최초)	2,844명
2	교육종단 연구	-	전국 중1 학 생(2005) 전국 중학교	교육통계연보 (학교명부)	다단계층화 군집추출법	학생, 학교, 학부모, 학교장, 교사	학생 6,908명 학부모, 학교 장, 교사 등
3	교육고용 패널	-	전국 중3, 고3 학생(2004)	교육통계연보 (전국 중·고등학교 학교명부)	층화집락 추출법	학생(중3, 고3), 보 호자, 학교(행정, 담임)	학생 6,000명, 보호자, 학교
4	청년패널	YP20 01	만 15~29세 청년층(2001)	95년 센서스 표본조 사구 - 산업직업별 고용조 사 20%	산업직업별 고용조사 에서 이중 추 출법	만 15~29세 청년	만 15~29세 8,296명(2001) 만 15~29세 5,956명(2003)
		YP20 07	만 15~29세 청년층(2007)	2006년 산업직업별 고용조사		만 15~29세 청년	만 15~29세 약 1만명 (2007)
5	대졸자 직업이동	-	전국 전문대 및 대학졸업자 (2004.8~2005.2)	고등교육기관 졸업자 취업통계조사(2005)	다단계층화 추출법	전국 전문대 및 대학졸업자 약 5% (2004.8~2005.2)	25,802명 (2006)
6	여성인력 패널	-	전국 100인 이상 기업의 대리급 2년차 이상 여성	한국신용평가 포함 100인 이상 기업	비례할당표집	대리급 2년차 이상 여성	2,500명
7	고령화 연구패널	-	만 45세 이상 중고령자 (1962년 이전 출생자)	인구주택총조사 (보통, 아파트 조사구)	집락층화표집	45세 이상 중고령자	10,254명
8	장애인 고용패널	-	15세 이상 범 정장애인	인구주택총조사 (표본조사) 제가 장애인 가구	층화계통추출	15세 이상 제가장 애인	5,000명
9	아동패널	-	2008년 출생 신생아	250명 이상 분만건수 (2006 상반기) 산부 인과의 3·4월 출생아 명부	27개 병원 유 중 무선 표 집 후 전수조사	신생아, 양육 부모	1,800명
10	인구패널	-	전국가구	전국가구명부 (인구주택총조사)	확률비례층화	미혼자, 기혼자, 청소년	12,000가구 (예정)
11	노동패널	-	도시거주 가구 및 15세 이상 가구원	인구주택총조사 (1995) 표본조사구	2단계 층화집락 계통추출	조사대상 가구의 15세 이상 가구원	5,000가구
12	여성가족 패널	-	전국 만19세 ~ 64세 여성	인구주택총조사 (2005) 표본조사구	다단계 층화 계통추출	표본가구내 가구원 중 만 19~64세 여성	8,500가구 (10,000명)
13	복지패널	-	전국 가구	인구주택총조사 (2005)에서 국민생활실태조사 (2006)	층화이중추출	표본가구내 가구원 중 15세 이상 가구원	7,072가구 (14,463명)
14	의료패널	-	전국가구	인구주택총조사 (2005)	층화추출	표본가구내	8,000가구(실 제 3,000가구)
15	노후보장 패널	-	50세 이상 가구 원이 있는 가구	인구주택총조사 (2000) 표본조사구	확률비례층화	50세 이상 가구원	5,133가구
16	사업체 패널	-	상용근로자 30 인 이상 사업 장과 공공기관	사업체기초통계조사 공공기관 (기획예산처 기준)	사업장: 층화 공공기관: 전수	인사담당자, 노무담 당자, 근로자 대표	4,275사업장 (일반 3,916, 공공 359)
17	인적자본 기업패널	-	근로자 100명 이상 일반기업	한국신용평가(2005)	산업, 규모, 기 업형태 고려할 당표집	기업, 사업장, 근 로자(팀장, 팀원)	450기업(205개 사업장), 근로자

대리급 2년차 이상인 여성을 모집단으로 하고, 한국신용평가에 포함된 100인 이상 기업을 표집틀로 활용하였다. 그런 다음 업종, 종업원 규모, 여성관리자 비중을 고려한 비례할당 표집법을 사용하여, 2,500명의 대리급 이상 여성관리자를 추출하였다. 아동패널의 경우 2008년 출생아인 모집단의 정보를 담고 있는 표집틀이 없어, 일정 조건(상반기 205명 출산)을 갖춘 산부인과 27개를 유층무선표집한 후에 조사 방식을 계획하고 있으나 병원 표집단계에서 엄격한 확률표집 원칙이 지켜지지 않고 있다. 인적자본 기업패널의 경우에도 산업, 규모, 기업형태를 고려한 할당표집법을 사용하고 있다.

5) 현장조사

가) 현장조사 진행

17개 패널 가운데 현장조사를 기관 자체에서 실시하는 경우는 한국보건사회연구원과 한국교육개발원뿐이다. 한국교육개발원의 경우에도 지금은 학교조직을 통해 현장조사를 진행하고 있지만, 향후 전문실사업체 외주를 고려하고 있다⁸⁾. 이 두 개 기관에서 운영하고 있는 4개 패널과 아직 1차 본 조사를 실시하지 않은 아동패널을 제외하면, 나머지 12개 패널 모두 전문실사업체에 의뢰하여 현장조사를 진행하고 있다. 외국의 경우를 보면 미국의 PSID는 자체조사, 캐나다의 SLID는 캐나다 통계청 지역사무소에서 현장조사를 진행하고, 독일의 GSEP와 영국의 BHPS는 민간업체에 위탁하고 있다 <표 2-6>. 우리의 경우 현장조사에 대한 전문실사업체 의존율이 지나치게 높다는 것을 알 수 있다. 한국보건사회연구원을 제외하면 자체 현장조사 조직을 보유하지 못한 연구소나 공단에서 패널조사를 운영하는 만큼 이러한 실사전문업체 의존은 불가피한 면이 없지 않으나, 민간업체의 현장조사 진행으로 인해 자료의 신뢰성 측면에 의문을 제기하는 경우도 없지 않다.

8) 교육중단연구의 경우, 학교단위 표집을 통해 조사를 진행한 관계로 제한된 중학교에 대해서만 조사를 진행하는 편이 있었으나, 향후 고등학교 진학단계에서 학생들이 흩어질 것으로 예상되기 때문에 향후 전문실사업체를 통한 현장조사 진행을 고려 중이다.

〈표 2-6〉 현장조사 진행

ID	조사명	운영기관	실사기관	실사기관 변경	조사방법	답례품
1	청소년패널	한국청소년 정책연구원	미 디 어 리 서치	없음	조사원면접 (학부모-전화)	1만원 상품권
2	교육종단연구	한국교육개발원	자체	없음	우편조사 (학교일괄-성취도)	학생, 학부모, 교사 (2천원 상당)
3	교육고용패널	한국직업능력 개발원	미 디 어 리 서치	없음	CAPI, CATI	상품권
4	청년패널	한국고용정보원	한국리서치 (YP2001)	변경 (자체-한 국-TNS)	조사원면접	상품권
			TNS (YP2007)		CAPI	
5	대졸자직업이동	한국고용정보원	한국리서치	비해당	조사원면접, CAPI(준비)	상품권(1만 5천)
6	여성인력패널	한국여성정책연구원 (여성부 수탁)	KDN	없음	조사원면접	미정
7	고령화연구패널	한국노동연구원	TNS	비해당	CAPI	현금(1~1.5만원)
8	장애인고용패널	한국장애인고용 촉진공단	한국갤럽	비해당	조사원면접 (예정)	-
9	아동패널	육아정책개발센터	비해당	비해당	비해당	비해당
10	인구패널	한국보건사회연구원 (보건복지가족부 수탁)	자체	없음	조사원면접	-
11	노동패널	한국노동연구원	한국리서치	없음	조사원면접, CAPI(시협)	물품(5천원)
12	여성가족패널	한국여성정책연구원	한국리서치	없음	CAPI	문화상품권
13	복지패널	한국보건사회연구원	자체	변경 (갤럽-자체)	조사원면접	상품권
14	의료패널	한국보건사회연구원	자체	없음	조사원면접	상품권
15	노후보장패널	국민연금관리공단 (국민연금연구원)	한국갤럽	없음	조사원면접	상품권 (가구당 3만원)
16	사업체패널	한국노동연구원	동서리서치	변경 (현대-동서)	CAPI	선물(넥타이)
17	인적자본기업 패널	한국직업능력개발원	KDN	없음	조사원면접	현금/상품권 /선물

현재 12개 패널의 외주 민간업체 현황을 보면 한국리서치(3개), 한국갤럽(2개), 미디어리서치(2개), KDN(2개), TNS(2개), 동서리서치(1개)이다.

패널조사는 그 특성상 조사기획에서부터 자료 분석 및 활용의 단계까지 일관된 관리가 필요하다. 현장조사에 대한 적절한 관리지침을 만들어 체계적이고 일관된 관리가 수반되어야 한다. 즉, 양질의 자료를 생산하기 위해서는 연구기관과 실사기관의 안정적이고 유기적인 관계가 필요하다. 그런데,

현장조사 기관 변경을 보면 길지 않은 조사 운영 기간 동안 청년패널과 사업체패널과 같이 업체를 변경한 경우도 있었다. 장기적으로 볼 때 담당 업체 변경은 자료의 질에 부정적 영향을 미칠 가능성이 있다.

현장조사의 자료수집 방법을 보면 전통적인 조사원 면접조사가 10개 패널에서 활용되고 있으나, CAPI를 이용하거나 도입을 위해 준비하는 패널이 점차 증가하고 있다. 이 외에도 학부모 대상 전화조사(청소년패널), 우편조사(교육종단연구), 학교대상 집단면접(교육종단연구), CATI(교육고용패널)를 이용하는 경우도 있다. 향후 자료입력 및 에디팅 프로그램의 활용이 증가하게 되면, 점차 전통적인 대인면접조사(PAPI)에서 벗어나 CAPI나 CATI 조사방법 활용이 증가할 것으로 예상된다. 이는 해외의 경우도 마찬가지로, PSID와 SLID는 CATI방식을 활용하고 있고, GSEP는 PAPI와 CAPI를 병행하고 BHPS는 CAPI를 활용하고 있다.

마지막으로 현장조사 응답 답례품은 대부분의 조사에서 제공하고 있으며, 상품권 형태가 가장 많고 선물이나 현금을 제공하는 경우도 있다. 외국의 경우 SLID는 면접원과 조사원의 관계를 중심으로 지원하되 현물보상은 하지 않으나, PSID, GSEP, BHPS는 모두 현금이나 상품권, 쿠폰 등을 제공하고 있다.

나) 패널 관리방법

패널조사의 성공을 위하여 가장 중요한 요인 중 하나는 표본의 탈락률을 최소화하는 것이다. 조사 차수가 증가함에 따라 탈락되는 표본이 특정 계층에 집중되면 표본의 대표성을 저해하는 요인으로 작용하게 된다. 따라서, 탈락을 최소화하기 위해 다양한 방법이 활용된다. 예를 들면, PSID는 인터뷰 시, 이사 후 주소 통보 시, 참여 거부자가 참여 유도 인터뷰 시에도 현금보상이 주어진다. 이 외에도 조사결과 요약 팸플렛 제공, 참여거부 가구에 대한 설득편지, 친구 및 친척의 이름 파악, 연구책임자가 면접원 애로사항 청취, 반송 가능 우편물 사용 등 다양한 방법이 활용된다.

아직 예비조사나 1차 본 조사만 실시하고 있는 패널을 제외하고, 11개 패널에서 조사응답자 관리를 위해 활용하고 있는 방안을 보면, ‘이사회

〈표 2-7〉 응답자 관리방법

조사명	소식지 제공	정기 연락	친척/동료 연락처	이사시 주소파악	생일 카드	면접원 수 기	분석결과 제공	온라인 커뮤니티	기 타
청소년패널	실시	실시	실시	실시	미실시	미실시	실시	실시	경품행사
교육종단 연구	미실시	미실시	미실시	실시	실시	미실시	실시	미실시	학교협력 교사 수당지급
교육고용 패널	미실시	미실시	실시	실시	미실시	미실시	미실시	실시	직업세계 책자제공
청년패널	실시	실시	실시	실시	실시	미실시	실시	실시	-
대졸자 직업이동	실시	실시	미실시	실시	실시	실시	비해당	실시	감사카드, 달력
고령화 연구패널	실시	실시	실시	실시	미실시	미실시	실시	미실시	달력제공
노동패널	실시	실시	미실시	실시	실시	실시	실시	미실시	협조공문, 감사편지, 감사선물
복지패널	실시	실시	실시	실시	미실시	미실시	실시	미실시	명절선물
노후보장 패널	실시	실시	미실시	실시	미실시	미실시	실시	실시	경품행사, 면접원 간담회, 감사편지
사업체패널	실시	실시	미실시	실시	미실시	미실시	실시	실시	-
인적자본 기업패널	실시	실시	미실시	실시	미실시	미실시	실시	미실시	-
실시/합계	9/11	9/11	5/11	11/11	4/11	2/11	9/10	6/11	

주소파악'은 모든 패널에서 실시하고 있고 '소식지 제공', '정기적 연락', '분석결과 제공' 등은 대부분의 패널에서 실시하고 있다<표 2-7>.

반면에 '평소 친척이나 직장동료 연락처 파악', '생일축하카드 발송'은 상대적으로 실시하지 않는 경우가 더 많았고, '온라인 커뮤니티'는 11개 패널 중 6개에서 제공하고 있다. 기타 응답자 관리 방법으로는 경품행사, 감사편지(카드), 달력제공, 명절선물 등이 응답되었다. 우리의 경우 전문실사업체를 통해 현장조사를 진행하는 비율이 높기 때문에, 응답자인 패널 관리업무도 해당 업체를 통해 진행될 가능성이 높다. 민간기업의 경우, 이윤추구로 인해 응답자 관리를 소홀하게 할 개연성이 항상 존재한다. 예컨대 연간 인구이동률이 20%에 달하는 우리 사회에서

이사 시를 대비해 친척 및 동료 연락처 파악은 필수적인데, 이를 이행하는 않는 패널도 적지 않다. 이는 결국 표본탈락의 가능성을 높혀 표본의 대표성에 영향을 미칠 수 있다.

다) 패널유지율

표본유지율을 알아보기 위해 1차조사 응답률, 표본유지율 및 항상응답률(always in)을 조사하였다. 표본유지율은 1차 원표본 대비 해당 차수 표본유지율을 의미하고, 항상응답률은 모든 조사차수에 빠지지 않고 응답한 응답자 비율을 의미한다. 따라서, 2차년에는 표본유지율과 항상응답률이 동일하지만 이후에는 유지율보다 항상응답률이 낮은 것이 일반적이다. 표본유지율은 조사단위에 따라 기준이 계산되며, 가구단위의 경우 모든 조사대상 가구가 응답하지 않아도 가구주가 조사되면 일반적으로 응답가구로 간주된다.

먼저, 개인 및 가구패널의 1차 조사 응답률⁹⁾을 보면 교육중단연구에서 99.6%, 대졸자 직업이동에서 62.2% 등 다양한 응답률을 보이고 있다. 이는 현장조사가 진행되는 상황에 따른 것으로 추측된다. 교육중단연구의 경우, 표집된 학교를 대상으로 해당 학교의 도움을 받아 진행되기 때문에 높은 응답률을 보이는 것이 당연하다. 1차 조사시 학교단위 조사를 실시한 청소년 패널(중 2패널)의 1차 응답률은 66.3%로 낮은 편이다. 대졸자 직업이동의 경우 사전접촉을 통해 조사협조를 약속한 표본대비 실제 응답률이 62.2%이다. 1차 응답률이 가장 낮은 조사는 노후보장패널이 42.7%로 매우 엄격한 표본선정 조사기준을 적용한 것에서 기인한 것이다. 나머지의 경우, 1차 조사에서 응답률을 산정할 수 없었거나 혹은 계산하지 않은 경우도 있으나 주로 70~80% 수준으로 해외 수준과 유사하다. 예를 들면, 고령화연구패널의 경우 1차 본 조사의 응답률이 70.7%(가구기준)이고, 개인 기준으로는 89.2%이다. 유사한 외국 패널의 경우 미국 HRS(Health and Retirement Survey, 1992)는 가구기준 80.2%, 개인기준 81.6%이고, 영국의 ELSA(2004)도 각각 69.9%, 96.5% 수준이다 (한국노동연구원 고령화연구패널조사팀, 2007).

9) 미국 PSID, 독일 GSEP, 영국 BHPS의 1차년도 응답률은 각각 76%, 65%, 70%이다(강석훈, 1998).

표본유지율에 앞서 알아둘 것은 3개의 패널, 즉 청년패널, 복지패널, 사업체패널이 1차 본 조사 후에 다시 원표본을 재구성했다는 사실이다. 청년패널의 경우 2001년에 본 조사를 시작했으나 2002년에 원표본을 재구성하여, 2차년도에 구축된 5,956명을 원표본으로 하여 추적조사를 실시하고 있다. 그리고, 사업체패널의 경우에도 WPS2002와 WPS2003을 동일 표본으로 조사한 후에 2005년에 표본을 개편하여 WPS2005 조사를 실시하였다. 또한 복지패널도 ‘차상위·빈곤층 패널사업’, ‘저소득층자활사업실태조사’, 서울대학교의 ‘한국복지실태조사’를 2006년에 통합한 것이다. 이와 같은 이유로 복지패널은 1차년도 본 조사가 2006년이고, 청년패널(YP2001)은 2001년과 2002년 두 가지 기준으로 유지율을 작성하였고, 사업체 패널의 경우 1차 응답률은 WPS2005년 기준이고, 2차년도는 WPS2003년을 기준으로 작성하였다.

본 조사가 2차 이상 진행된 9개 패널을 대상으로 패널유지율과 항상 응답률을 보면, 외국과 마찬가지로 개인패널의 경우 패널유지율이 전체적으로 높고, 가구패널은 상대적으로 낮은 편이다¹⁰⁾. 청소년패널과 교육종단연구의 경우 학교단위 표집을 하고 있어 2차 표본유지율이 90%를 상회하고 있고, 또한 교육고용패널의 고등학교 3년 패널의 경우 군입대로 인해 조사가 불가능한 표본이 3차 이후 급격히 증가해 3차년도 표본유지율이 68.2%로 급속히 낮아지고 있다. 하지만 조사불능 사례를 제외한 유효표본의 유지율은 2차 97.0%, 3차 75.4%이다. 대졸자직업이동의 2차 표본유지율은 88.9%로 높은 편이다. 다만, 청년패널의 경우 2002년을 기준으로 하면 표본유지율이 3차 85.8%, 6차 75.8%로 큰 문제가 없으나 2001년 원표본을 기준으로 하면 표본탈락이 심각한 수준이다. 그리고 모든 차수에 응답한 비율이 5차 년도에 이미 54.1%로 표본의 대표성 문제를 안고 있는 것으로 판단된다. 한편, 가구패널의 경우 패널유지율이 이미 국제적 수준으로 평가되는 노동패널을 제외하면 3차 이상의 표본유지율에 대한 정보가 없어 아직 평가하기는 이르다. 10차까지 조사가 진행된 노동패널의 경우 외국의 가구패널과 유사하게 3차년도 이후 70%의 안정적인 패널유지율을 보이고 있다. 다만, 항상응답

10) 가구단위 조사인 미국의 PSID의 경우 10차 표본유지율이 70%이고, 개인단위 조사인 NLSY의 10차 표본유지율은 90.2%로 개인표본의 유지율이 뚜렷이 높다(신동균, 1998).

34 농가소득의 동태적 변화요인

률이 9차년도에 55.7%로 1차~9차 모두에 응답한 가구가 원표본의 과반을 상회하는 수준이다. 이에 따라 앞으로 표본 확대나 개편에 대한 연구가 진행되어야 할 것으로 보인다. 하지만 현재 전체적인 표본유지율이 높은 수준이 아니고, 또한 패널관리에 대한 노력이 철저하지 못해 향후 패널 유지율은 지속적인 관심을 가지고 지켜봐야 할 과제이다.

<표 2-8> 표본유지율 및 항상응답률(always in)

(단위 : %)

조사명	복 수	1차 응답률	구 분	2차	3차	4차	5차	6차	7차	8차	9차
청소년 패널	중2 패널	66.3	표본유지율	92.4	90.6	90.5					
			항상응답률	92.4	87.5	84.4					
	초4 패널	81.0	표본유지율	95.2	94.0						
			항상응답률	95.2	91.7						
교육종단 연구	-	99.6	표본유지율	99.3	98.8						
			항상응답률	97.1							
교육고용 패널	-	무응답	표본유지율	87.6	68.2						
			항상응답률 (유효기준)	90.4	65.3						
청년패널	YP2001 기준	74.9	표본유지율	71.8 (2001기준) 85.8 (2차 기준)	61.6	60.3	57.5	54.4			
			항상응답률	82.1	70.1	59.4	54.1				
대출자 직업이동	-	62.2	표본유지율	88.9							
			항상응답률	88.9							
노동패널	-	75.5	표본유지율	87.6	80.9	77.3	76.0	77.2	77.3	76.5	76.5
			항상응답률	87.6	77.9	72.3	65.5	61.7	59.3	57.3	55.7
복지패널	-	무응답 (미산정)	표본유지율	92.1							
			항상응답률	92.1							
노후보장 패널	-	42.7	표본유지율	89.0							
			항상응답률	89.0							
사업체 패널	-	53.6 (2005)	표본유지율	80.1 (2003)							
			항상응답률	80.1							

주: 조사진행 차수가 2차 미만인 패널은 분석에서 제외함

6) 패널자료 활용

가) 조사결과 활용

본격적으로 중단분석을 실시할 수 있는 4차 이상 조사가 진행된 패널이 노동패널, 청소년패널, 청년패널, 교육고용패널이다. 자료생산 후 실제 분석까지 소요되는 기간을 고려하면 아직은 본격적인 연구결과를 기대하기 어려운 초기 패널이 대부분이다. 여하튼 생산된 패널자료를 활용해서 진행된 연구결과물을 보면 패널조사의 차수(wave)에 정비례해서 논문 및 보고서 산출물이 증가하는 것을 알 수 있다. 조사 차수가 10차인 노동패널의 경우 논문 415편이 나왔고, 사업체패널(6차)도 논문 87편, 보고서가 23권이고, 청소년패널(5차)은 논문 200편, 보고서가 12권이 나왔다. 하지만, 패널 조사자료의 본래적 특징인 중단분석을 실시한 사례는 청년패널의 경우만 13편으로 응답했고, 사업체 패널은 논문 87건 중에서 중단분석을 실시한 논문은 1편이라고 응답했다. 다른 패널에서는 파악하지 않고 있거나 응답을 기피하였다.

또한 패널조사의 분석결과에 근거해서 정부정책에 영향을 미친 사례나 다른 조사설계와는 달리 패널조사로 설계함으로써 관련분야에 기여한 내용을 질문하였다. 왜냐하면, 모든 패널조사들이 출발 시점에서 패널설계를 통해 정책 과제에 대한 평가나 점검이 가능할 것으로 보고 있고, 횡단면 조사나 횡단면 반복조사에 비해 관련 분야에 기여할 수 있을 것으로 예측하고 있기 때문이다. 청소년패널은 조사내용 전반에 걸쳐 종단적 변화실태 파악이 가능하고 그 원인 진단이 가능해졌으며, 나아가 정책과 관련해서는 청소년 기본계획 수립 및 조사내용이 참고자료로 활용되고 있다고 응답했다. 청년패널은 패널설계를 통해 청년의 미취업 탈출시기, 노동시장 진입 소요시간, 첫 직장 등에 대한 측정이 가능해졌다고 답했다. 보다 실질적으로 여성가족패널의 경우 예비조사 결과를 활용하여 ‘여성비정규직 모성보호 현안 정책’에 기여했고, 사업체패널의 경우에도 「육아휴직실태조사」로 육아휴직제도 개선과 「비정규직 수요조사」 결과를 활용하여 비정규직 관련 입법에 활용한 것으로 보고하였다. 사업체패널은 또한 향후 비정규직법 관련 영향에 대한 분석을 통해 관련 분야에 기여할 수 있을 것으로 예상했다. 가장 오랜 역사를

지난 노동패널의 경우 응답은 하지 않았지만, 노동 관련 분야 정책 결정에 다양한 참고자료로 활용되었다는 사실은 분명해 보인다. 패널조사의 역사가 짧은 점을 고려할 때, 아직은 관련 분야 기여도나 정책적 활용도가 그리 높은 편은 아니다. 하지만, 외국의 경우를 참고할 때, 향후 패널조사의 분석결과는 관련 분야 연구나 정책적 활용의 잠재력이 충분한 것으로 여겨진다.

한 가지 아쉬운 점은, 패널연구자들이 자료생산 및 분석에 대한 관심도는 높으나, 상대적으로 분석결과의 기여나 정책적 활용에 대한 관심이 저조하다는 사실이다. 정책적 활용 측면은 연구기관의 관심 범위를 넘어서는 영역으로 이는 관련분야 학자나 정책입안자의 몫으로 간주하고, 연구기관은 자료의 생산 및 보고서 작성을 주된 업무로 간주하는 경향이 있었다. 앞으로는 자료 활용에 대한 관심을 기울일 필요가 있을 것이다. 자료 활용과 관련해서 마지막으로 지적하고 싶은 것은 향후 패널자료를 이용한 중단분석 기법에 대한 체계적인 지원이나 교육이 필요하다는 것이다. 국내에서 패널자료를 활용가능하게 된 지가 오래되지 않아 아직 패널자료를 적절하게 생성하고 다양한 중단분석을 할 수 있는 저변이 취약한 것으로 여겨진다. 이로 인해 양질의 자료를 적절하게 분석, 활용할 수 있는 능력이 아직은 취약해, 기존의 횡단적 분석을 반복하는 경향이 있다. 이 문제를 해결하기 위해, 패널자료 생성 및 분석기법에 대한 교육기반이 확충되어야 할 것이다.

〈표 2-9〉 패널조사의 기여점 및 정부정책 활용사례

조사명	패널조사의 기여점	정부정책 활용사례
청소년패널	조사영역(일, 진로, 비행, 여가, 사회적 관계 등) 관련 중단적 변화실태 및 원인 진단 추적	청소년 기본계획 수립시 참고, 조사영역(일, 진로, 비행, 여가, 사회적 관계 등) 관련 참고자료 활용
청년패널	미취업 탈출시기, 노동시장 진입 소요시간, 첫직장 근속기간 측정으로 이동경로 추적가능	무응답
여성가족패널	비해당	예비조사 결과로 여성비정규직의 모성보호 현안 정책에 기여
사업체패널	비정규직법 관련 영향에 대한 분석이 2008년에 진행 예정	WPS2002의 육아휴직실태조사는 육아휴직제도 개선에 기여, WPS2002, 2003의 비정규직 수요조사는 비정규직 관련 입법에 활용

나) 자료이용

정부 예산을 투입하여 생성된 패널자료는 관련 연구분야 종사자들이 다양하게 분석, 활용할 수 있어야 한다. 이를 위해서는 자료의 생산단계에서 잠재적 자료 이용자의 요구를 반영하고, 원시자료를 손쉽게 이용할 수 있도록 해야 한다.

먼저, 이용자 요구 및 의견 수렴 방법을 알아본 결과, 조사결과 및 자료 이용자 집단의 요구를 반영하는 경로로 대부분의 패널에서 전문가 자문회의, 학술대회를 활용하고 있는 것으로 나타났다. 자료 이용자에 대한 만족도를 조사하는 경우는 한국고용정보원의 청년패널과 대졸자 직업이동으로 2개 패널에 그치고 있다.

그리고 원시자료의 제공방법을 보면, 기관에 따라 차이가 있으나 홈페이지를 통해 무료로 제공하는 경우가 점차 증가하고 있으며, CD 형태로 유료판매하거나 혹은 학술대회를 통해서만 제공하는 경우도 있다. 예외적으로 한국교육개발원은 유료판매 혹은 신청 후 기관 동의 하에 자료를 제공하고 있다. 자료 활용의 활성화와 원시자료의 공공재적 성격을 고려하면, 향후 이용자들이 손쉽게 무료로 이용할 수 있도록 이용자가이드와 함께 무료 제공하는 것이 바람직하다.

〈표 2-10〉 자료활용 및 제공

ID	조사명	차수	논문	보고서	중단분석	원자료 제공	이용자요구반영	
1	청소년패널	5	200	12	미과악	학술대회 참가자 CD판매	자문회의/학술대회/ 데이터이용자의견	
2	교육중단연구	3	34	2	비해당	자료활용 계획서 제출 및 동의서 신청 후 제공	학술대회/전문가협의/ 세미나	
3	교육고용패널	4	68	5	미과악	무료제공 (웹사이트)	학술대회/전문가협의/ 세미나	
4	청년패널	7	62	10	13		고객만족 설문조사	
5	대졸자직업이동	2	무응답	무응답	무응답	무료제공 (웹사이트)	전문가의견수렴/ 만족도 조사	
6	여성인력패널	1	비해당					
7	고령화연구패널	2	0	무응답	0	홈페이지	전화/이메일	
8	장애인고용패널	0	비해당					수요조사/학술대회/간담회
9	아동패널	0	비해당					
10	인구패널	0	비해당					
11	노동패널	10	415	미집계	미집계	무료제공 (홈페이지)	학술대회	
12	여성가족패널	1		1			전문가의견수렴/ 자문기획단회의	
13	복지패널	2	6	3	비해당	무료제공 (홈페이지)	홈페이지	
14	의료패널	1	비해당					
15	노후보장패널	3	무응답	무응답	무응답	학술대회	전문가자문회의/학술대회	
16	사업체패널	6(3)	87	23	1	홈페이지 제공 (유상/무상)	홈페이지/학술대회	
17	인적자본기업패널	2	28	4	비해당	CD유료판매	학술대회/전문가 회의/ 워크샵	

7) 예산 및 인력

가) 패널 예산

마지막으로 17개 패널 운영과 관련된 예산 및 인력 현황에 대해서 살펴보고자 한다. 장기적인 중단연구를 수행하기 위해서는 안정적인 예산과 적정 인력 확보가 필수적이다. 먼저 예산에 대한 조사결과를 보면, 패널 전체예산(사업비)은 2000년 이후 2년마다 2배 이상 증가하는 급속한 증가세를 보이고 있다. 1998년 438백만원, 2000년 454백만원, 2002년 948백만원, 2004년 2,579백만원, 2006년 7,368백만원, 2007년 패널예산은

11,158백만원이다. 2007년을 기준으로 예비조사 단계인 ‘아동패널’과 ‘인구패널’ 예산을 제외한 10,828백만원을 15개 패널로 나뉘보면, 본 조사가 진행 중인 15개 패널의 평균 예산은 약 722백만원 수준이다. 더욱이 패널예산은 사업비만을 의미하므로 패널 운영기관 연구인력의 인건비까지 계산하면 현재 사업예산보다 훨씬 더 많은 예산이 투입되고 있는 것이다. 이와 같은 예산증가의 원인은 기존에 운영 중인 패널 예산의 지속적 증가와 함께 신규 패널이 생성되고 있기 때문이다. ‘청소년패널’과 ‘교육중단연구’ 패널을 제외하면 운영 중인 모든 패널사업의 예산이 증가하고 있다. 예를 들면 ‘노동패널’의 경우 1998년 438백만원에서 2007년 850백만원으로 두배 가량 증가했다. 이와 같이 기존 운영 패널의 경우에도 통상적인 예산증가와 함께 복수의 패널운영 등을 통해 계속적으로 사업비가 증가하고 있으며, 또한 지속적인 신규 패널의 출발로 예산이 가파르게 상승하고 있는 것이다. 2008년에도 ‘장애인고용패널’과 ‘아동패널’의 본 조사가 예정되어 있고, 여러 정부부처 및 연구소에서 추가적인 패널운영 계획을 시도하고 있어 앞으로도 이러한 예산 증가세는 지속될 것으로 예상된다.

패널사업에 투입되는 예산의 성격에 따라 구분해보면 일반예산이 11개 사업, 기금예산이 6개 사업이다. 기금사업으로는 고용보험기금을 사용하는 ‘청년패널’, ‘대졸자직업이동’, ‘고령화연구패널’, ‘노동패널’과 국민연금을 활용하는 ‘노후보장패널’, 장애인고용촉진 및 직업재활 기금의 ‘장애인고용패널’이 있다. 예산 성격에 따른 패널 당 평균예산을 보면 기금이 904백만원, 일반예산이 568백만원(아동패널은 예비단계로 제외)으로 기금예산으로 운영하는 패널의 예산이 더 큰 것으로 나타났다. 예산규모만을 보면 ‘대졸자직업이동’의 경우 대규모 표본을 구성하면서 2007년 예산이 1,761백만원으로 가장 많은 예산을 사용하였고, 같은 시기에 ‘복지패널’은 1,452백만원으로 그 뒤를 잇고 있다.

40 농가소득의 동태적 변화요인

<표 2-11> 패널조사 예산

(단위 : 백만원)

조사명	예산구분	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	비고
청소년 패널	일반회계	-	-	-	-	-	150	500	550	470	439	
교육 중단연구	일반회계	-	-	-	-	-	-	55	320	279	250	
교육고용 패널	일반회계	-	-	-	-	-	500	500	500	500	800	
청년패널	기금 (고용보험기금)	-	-	-	100	310	320	321	330	400	779	
대졸자 직업이동	기금 (고용보험기금)	-	-	-	-	-	-	-	-	1,854	1,761	
여성인력 패널	일반회계	-	-	-	-	-	-	-	-	-	238	돌봄 종사자 조사 포함
고령화 연구패널	기금 (고용보험기금)	-	-	-	-	-	-	-	760	769	769	
장애인 고용패널	기금(장애인고 용촉진 및 직 업재활기금)	-	-	-	-	-	-	-	-	-	570	
아동패널	일반회계	-	-	-	-	-	-	-	-	-	50	
인구패널	일반회계	-	-	-	-	-	-	-	-	-	280	
노동패널	기금 (고용보험기금)	438	390	454	609	638	650	673	688	846	850	2005년 이전 예산은 현장 조사비용
여성가족 패널	일반회계	-	-	-	-	-	-	-	-	150	770	
복지패널	일반회계	-	-	-	-	-	-	-	-	1,300	1,452	
의료패널	일반회계	-	-	-	-	-	-	-	-	-	500	
노후보장 패널	기금(국민연금)	-	-	-	-	-	-	-	570	-	700	
사업체 패널	일반회계	-	-	-	-	-	180	230	-	500	500	
인적자본 기업패널	일반회계	-	-	-	-	-	-	300	300	300	450	
합계		438	390	454	709	948	1,800	2,579	4,018	7,368	11,158	

주: 1) 예산은 연구인력 인건비를 제외한 사업비를 응답한 것임
 2) 예산구분은 최근 예산의 성격으로 구분한 것임

나) 패널 인력현황

한편, 패널사업에 관계하고 있는 연구기관의 인력은 17개 사업에 총 88명으로 단위 패널 당 평균 5.2명(최소 3명~최대 10명) 수준이다. 연구

인력의 학력을 보면 박사가 36명으로 전체 연구인력의 41%를 차지하여 패널 당 2명(2.1명) 가량이 투입되고 있으며, 석사 이하 연구인력은 전체 패널에서 52명인 59%로 패널 당 3명(3.1명) 가량이다. 연구 인력 교체 등을 고려하면 패널의 안정적 운영을 위해서는 최소 박사급 인력을 두 명 이상 확보할 필요가 있다.

이들을 고용형태별로 나뉘보면 정규직이 65명(74%), 비정규직이 23명(26%)으로, 패널 당 평균 정규인력 3.8명, 비정규인력 1.4명으로 연구진이 구성되어 있다. 패널사업의 경우 장기간의 안정적 운영이 중요하므로 비정규인력을 최소화할 필요가 있다. 단일 패널운영팀인력의 정규직 비율이 70%에 미치지 못하거나 3명 이하인 경우는 사업의 안정적 운영에 위협요인이 될 수 있을 것으로 여겨진다.

그나마 많지 않은 정규직 인력의 업무참여도를 보면, 문제의 심각성을 알 수 있다. 노후보장패널, 장애인고용패널과 같이 공단에서 운영하는 패널의 경우 100%로 패널업무만을 담당하고 있으나, 정부출연연구기관의 경우 한국노동연구원에서 운영하는 3개 패널을 제외하면 전반적으로 패널업무 외에 다른 연구업무와 병행하고 있는 것을 알 수 있다. 특히 정규인력의 패널업무 참여도가 50%이하인 패널은 9개로 모두 정부출연연구기관에 집중되어 있다. 이들 연구진들은 패널사업 외의 다른 연구업무에 투입되고 있는 것으로 판단된다. 이들 연구기관의 경우 한국노동연구원과 같이 패널전담 사업팀을 꾸려 패널운영을 추진해야 할 것이다.

연구자들의 패널조사 관련 경력을 보면, 연구에서 주도적 역할을 담당하는 박사인력의 평균 경력이 1년 이하인 경우가 5개 패널(대졸자직업이동, 여성인력패널, 장애인고용패널, 복지패널, 사업체 패널), 2년 이하가 3개 패널(아동패널, 여성가족패널, 의료패널)로 나타나, 패널에 대한 전문적인 지식이나 경험이 부족한 인력에 의해 패널이 운영되는 경우가 다수를 차지하고 있다. 패널조사의 역사가 길지 않고 단시간에 급속히 확산된 사정을 고려해 볼 때, 패널을 운영할 수 있는 전문적인 인력부족은 패널자료의 질과 결과 활용에 심각한 문제를 야기할 수 있다. 장기적인 종단연구인 패널조사의 특성을 고려할 때, 연구책임자의 잦은 교체 또한 문제이다. 전체 패널에서 연구책임자 교체가 있었던 패널이

42 농가소득의 동태적 변화요인

8개나 되고 청년패널, 노후보장패널, 노동패널은 3번 이상 연구책임자를 교체한 경험이 있다. 이와 같이 패널운영 경험 부족과 연구책임자의 잦은 교체로 인해 많은 예산이 투입되고 있음에도 불구하고 자료의 질에 부정적 영향을 미칠 가능성은 충분하다.

<표 2-12> 패널별 인력현황

ID	조 사 명	총인원	박사	석사 이하	정규직 (%)	비정규직	정규직 업무 참여도	박사 경력 평균	석사 경력 평균	책임자 교체
1	청소년패널	6명	3명	3명	3명(50%)	3명	70%	48월	8월	없음
2	교육 중단연구	4명	2명	2명	2명(50%)	2명	50%	35.5월	28월	2번
3	교육고용패널	7명	2명	5명	4명(57%)	3명	38%	41월	38.2월	1번
4	청년패널	3명	1명	2명	3명(100%)	0명	73%	72월	36월	3번 이상
5	대졸자 직업이동	4명	2명	2명	4명(100%)	0명	65%	8.5월	44.5월	없음
6	여성인력패널	4명	3명	1명	4명(100%)	0명	25%	8월	0월	없음
7	고령화연구패널	4명	1명	3명	4명(100%)	0명	95%	58월	30월	없음
8	장애인 고용패널	4명	2명	2명	4명(100%)	0명	100%	6월	7월	없음
9	아동패널	3명	2명	1명	3명(100%)	0명	25%	24월	12월	없음
10	인구패널	5명	3명	2명	5명(100%)	0명	25%	28월	12월	없음
11	노동패널	6명	1명	5명	6명(100%)	0명	92%	43월	40.8월	3번 이상
12	여성가족패널	4명	2명	2명	3명(75%)	1명	30%	24월	45월	1번
13	복지패널	10명	2명	8명	5명(50%)	5명	50%	12월	12월	없음
14	의료패널	7명	4명	3명	5명(71%)	2명	50%	13월	10월	없음
15	노후보장패널	5명	2명	3명	2명(40%)	3명	100%	49월	0월	3번 이상
16	사업체패널	4명	1명	3명	4명(100%)	0명	83%	12월	21월	2번
17	인적자본 기업패널	8명	3명	5명	4명(50%)	4명	36%	31월	24월	1번
	합계 (평균)	88명 (5.2명)	36명 (2.1명)	52 (3.1명)	65명 (3.8명)	23명 (1.4명)	-	-	-	-

주 : 외부 공동 혹은 객원 연구원은 제외한 운영기관 내부 인력임.

다. 국외 패널조사

1) 미국 : The Pannel Study of Income Dynamics(PSID)

가) PSID 개요

존슨(Johnson) 대통령 재직시 빈곤에 대한 전쟁의 일환으로 Office of Economic Opportunity(OEO)는 통계국으로 하여금 이 빈곤에 대한 전쟁이 국민의 경제적 복지수준을 얼마나 향상시켰는지를 평가하게 되었다. 소위 'Survey of Economic Opportunity'라고 불리는 이 연구는 1966년에 3만 가구를 조사하였으며 이 조사는 1967년에도 진행되었다. 이러한 국민복지에 대한 지속적인 관심으로 OEO는 미시간대학 부설 Survey Research Center(SRC)로 하여금 그들이 보유하고 있던 SEO 표본으로부터 저소득 계층에 속하는 약 2,000가구를 추출하여 계속 조사할 것을 권유하였다. 이에 SRC 소속의 모건(Morgan) 교수는 이 연구가 중산층 및 부유층까지도 포함하여 전체 미국인들을 대표하는 연구가 되게 하기 위하여 SRC가 가지고 있는 표본조사 틀로부터 새로운 대표표본을 추출하여 기존의 SEO 표본에 추가시켰고 이 연구의 이름이 PSID이다.¹¹⁾

PSID만이 가지고 있는 대표적인 2가지 특징은 개개인들을 가구라는 틀속에서 오랜 세월 관찰해 왔으며, 가구들을 세대(generation)를 넘어서까지 추적하였고 종종 한가구로부터 파생된 다수의 세대들을 동시에 면접하였다는 것이다. 이러한 특징들은 구체적으로 다음과 같은 장점들을 가지게 한다.

① 어느 시점에서든 모든 연령층에 대하여 미국의 가구와 개인들을

11) 패널연구에서는 표본들이 1차년도 뿐만 아니라 해를 거듭하여도 계속 대표성을 갖게 하는 작업이 매우 중요하다. 패널표본이 어느 시점에서든 항상 모집단을 대표하기 위한 이상적인 조건은 첫째, 1차년도에 임의적으로 추출된 표본가구를 빠짐없이 조사해야 하며(100%의 응답률), 둘째, 최초 표본가구들이 하나도 빠짐없이 매년 조사되어야 하며(0%의 이탈률), 셋째, 표본가구들이 시간의 경과에 따른 모집단의 변화를 100% 모방할 수 있도록 해야 한다. 세 번째에 관련된 예를 들면 출생, 사망 등에 의한 인구의 변화, 결혼, 이혼 등에 의한 가구수의 변화 등이 표본에 의해서도 그대로 포착되도록 표본관리 및 추적원칙을 적절히 설정해야 한다. 이를 가장 명확하게 실행한 연구 중의 하나가 PSID이다.

대표하며, ② 고용, 소득, 인구학적인 정보에 대한 연도별(때로는 월별) 시계열 데이터를 장기간 보유하고 있으며, ③ 각 세대로부터 직접 얻는 성인들에 대한 장기간의 시계열 자료, 유년기뿐만 아니라 그들이 성년이 된 후에도 획득한 정보, 가구내의 모든 어린이들을 비교분석할 수 있는 자세한 정보들을 포함하여 세대간을 걸쳐 광범위한 정보를 가지고 있으며, ④ 건강, 부, 저축, 친족관계 등의 다양한 보충적인 주제들을 포함시켰으며, ⑤ 최근에는 이웃환경, 건강과 사망 등에 대한 특수목적의 연구주제들을 포함시켰으며, ⑥ 연구시작연도인 1968년도부터 종·횡단면적인 데이터를 계속 검토하고 있다.

이러한 장점들을 가지고 있는 PSID는 인구학자, 사회학자, 심리학자, 경제학자 등 여러 분야의 학자들의 관심을 끌었으며 대표적인 연구주제들은 소득분포의 동태적 변화, 소득이전사업, 저축과 부, 무보수 생산활동, 유년기 및 성년기에서의 빈곤과 사회보장, 친족망, 10대 임신부, 세대간부의 변화과정 등이다. 이처럼 PSID가 다양한 연구주제를 수행할 수 있었던 것은 좁은 주제에 대한 광범위한 질문이 아닌 광범위한 주제에 대해 적은 양의 설문을 수행했기 때문이다.

나) PSID 표본

PSID의 최초 표본은 2개의 독립적인 표본들로 구성되어 있다. 하나는 군대와 감옥, 정신병원 등을 제외한 생산가능 인구에 대한 횡단면적이 대표표본을 다단계층화추출 방법에 근거하여 추출한 표본이며, 다른 하나는 저소득층에 대한 표본이다.¹²⁾ 이 횡단면적인 표본은 SRC에 의해 추출되었고 종종 SRC 표본이라 불린다. 이는 등확률표본으로 1968년에는 2,930가구들을 성공적으로 면접하였다. 저소득층에 대한 표본은 PSID가 SEO 표본으로부터 추출한 1,872가구패널로 구성되어 있다. 이는 종종 SEO 표본이라 불린다. SEO 표본은 표준도시통계지역(standard metropolitan statistical areas: SMSAs)과 남부지역의 비표준도시통계지역

12) PSID는 1968년도 최초 표본 구성원의 25세 이하의 자녀 및 양자, 양녀로서 군대, 대학 기숙사 등의 기관에 속해 있는 사람들도 표본 구성원으로 분류하였다. 그들은 그 기관을 떠나 자신들의 가정을 구성한 후에는 면접되었다.

(non-SMSAs)으로 그 범위가 제한되어 있고, 이는 확률 표본이다. SRC와 SEO 모두 1차년도에 무응답자들이 있었다. SEO 표본의 무응답 이유는 다음과 같다.

첫째, SEO 표본을 추출하였던 원래의 센서스 조사에서부터 무응답이 있었다.

둘째, 추출된 센서스 표본 구성원들은 통계당국에 의해 그들의 이름이 SRC에 공개되어도 좋다고 권유받았으나 약 25%의 인원이 서명하기를 거부하였다.

셋째, OEO는 추출된 표본 구성원들 중 일부의 주소를 SRC에 보내지 못했다.

SRC와 SEO 표본 모두 확률 표본이므로 그들의 조합 또한 확률 표본이다. 그러나 그 조합은 등확률 표본이 아니다. 따라서 적어도 기술적인 통계치들을 사용함에 있어서는 적절한 가중치들을 사용해야 한다. 가중치를 사용해야 할 또 하나의 필요성은 1차년도와 그 이후의 조사들에서 무응답률이 다르다는 데에도 있다. 다시 말해 PSID 데이터에 적용되는 가중치들을 1차년도에 있어서의 상이한 선택확률(unequal selection probabilities)과 매년 달라지는 표본이탈률(differential attrition rates)을 보상하기 위함이다.

2003년도 표본선정을 보면 이전의 조사와 달라진 것은 없고 추가적인 표본선정도 없었다. 2001년 이후 표본들은 주표본집단과 신표본집단(이민자집단)의 두 부분으로 구분되었다.

주 표본집단(Core Sample)은 1968년에 선정된 최초 표본가구와 그들로부터 파생된 가구들이다. 이 집단에서 2001년에 성공적으로 면접조사를 마친 가구는 6,915가구이고, 1999년 또는 2001년에 면접조사에 응하지 않아 재조사(재접촉)를 해야 하는 가구는 340가구였다. 이들 7,255가구는 주 표본집단으로서 2003년 면접조사 대상이 되었다. 면접조사가 진행되는 동안 주 표본집단으로부터 분리된 가구들의 새로운 샘플로 추가되었다. 이렇게 추가된 표본은 792가구로 이들은 ‘조사완료집단’(Spilt-offs)으로 구분되었다. 이 파생집단은 6,915가구 중에서 파생된 가구이다. 한편 또 다른 추가된 표본은 재조사를 실시해야 하는 가구에

서 파생된 33가구들로 이들은 '재조사 집단'(Recontacts or Spilt-Recons)으로 구분되었다.

두 번째 표본집단은 대개 '신표본가구(New Sample)'로 불리는데 조금 더 정확하게는 '이민자집단(Immigrant Refresher Sample or Immigrant Sample)'을 말한다. 이 집단은 1968년 후반 이민으로 인해 미국 내 인구 구성의 변화가 생기면서 그 당시 이민자들을 1997년과 1999년 조사부터 패널조사에 포함시켰고, 1997년부터 2001년까지의 조사에서는 그들로부터 파생되어 새롭게 생겨난 가구들에 대해서도 조사하였다. 이들은 총 491가구로 2003년 조사에서도 면접대상에 포함되었다.

2003년도 응답률을 보면 93.5%로 총 7,870가구에 대한 면접이 성공적으로 이루어졌다. 주표본집단은 6,554가구로 이중 Spilt-offs 가구는 561가구였고 83.1%의 응답률을 보였다. 또한 주표본집단 중 재접촉으로 분리되었던 가구 중 333가구가 조사에 응하였고 64.6%인 215가구가 면접을 마쳤다. 한편 이민자 표본에서는 459가구(94.3%)가구가 면접조사에 응했고, 36개의 Spilt-offs 가구에서 61%의 응답률을 나타냈다. 최초의 이민자 표본에서 파생된 새로운 이민자 표본 93가구 중 45가구에서 50.6%의 응답률을 보였다.

최초 PSID 가구들은 1967년도의 미국 가구들을 대표하는 확률표본을 구성한다. 어느 시점에서든 또한 시간의 경과에 따라서도 이 표본이 대표성을 갖게 하기 위해 PSID는 다음의 표본추적원칙을 세웠다.

기본원칙은 1차년도의 표본가구들에 나타난 18세 이상의 가구(다시 말해 군대, 감옥, 기숙사 등의 가구가 아닌 기관들을 제외한) 구성원들을 PSID '표본 구성원'들이라 칭하고 이들을 추적 조사한다. 이 최초의 '표본구성원'이 이혼을 했을 경우에도 전남편과 아내 모두 '표본구성원'으로 남아 있으며, '표본구성원'들로부터 태어난 아이들도 '표본구성원'으로 분류된다. 또한 외국으로 이주하여도 '표본구성원'으로 남는다. 이들 중 면접대상은 18세 이상의 '표본구성원'이 가구에 있지 않고 군대, 감옥, 기숙사 등의 기관에 들어갈 경우는 이 사실을 그들이 떠난 가구에 기록 보존한다. 그 후 기관에서 나와 원래의 가구로 돌아가거나 새로운 가구를 형성할 때에는 다시 면접을 시작한다.

PSID는 표본구성원들 뿐만 아니라 그들과 같은 가구 내에 동거하고

있는 모든 사람들(편의상 동거인이라 부름)도 조사한다. 만약 ‘표본구성원’들이 이혼 및 재혼을 통하여 별도의 가구를 형성할 때에는 이들을 추적조사하거나 ‘동거인’들이 ‘표본구성원’들과 분리되어 별도의 가구를 형성하면 더 이상 추적조사하지 않는다.

이러한 과정은 모집단에서의 가구변화과정을 모방하여 표본이 시간을 초월하여 동태적으로 모집단을 대표할 수 있게 하기 위함이다. 최초의 표본가구들로부터 생겨난 아이들이 성장하여 새로운 PSID 가구들을 형성하고 최초의 PSID 가구들이 이혼 및 재혼을 통하여 별도의 PSID 가구들을 만들어냄에 따라 가구수와 개인 수는 계속 증가하였다.

PSID는 매년 면접조사를 통하여 개인들과 그들이 속해 있는 가구들에 대한 정보를 수집한다. 가구를 대표하는 성인 1명(주로 남성가구주¹³⁾)이 응답자가 된다. 때로는 어떠한 이유 없이 남편이 응답할 수 없는 경우에는 아내(혹은 동거인, PSID에서는 이를 ‘아내’라고 칭함)가 면접에 응하기도 한다. 이 가구주 혼자서 자신과 아내와 모든 가구 구성원들에 대한 정보를 제공한다. 가구내의 모든 구성원들에 대한 정보를 획득하고 있으나 가구주에 대해서는 보다 세밀한 정보를 수집하였다. 하지만 1970년 후반부터는 아내들에 대해서도 같은 정도의 자세한 정보를 수집하였다.

다) PSID 조사

PSID 데이터 수집은 미시간에 있는 SRC의 실사 사무국에 의해 수행된다. 여기에는 SRC의 실사감독관 및 SRC 관리자뿐만 아니라 PSID 참모들도 참여한다. 우선 실사전의 준비물로 설문지, 표지 및 면접원들을 위한 지침서 등이 있으며 이것들은 매년 갱신되어진다.

데이터 수집은 매년 3~9월까지 진행된다면 이는 SRC의 실사 사무국이 전국에 흩어져 있는 실사 감독관들 및 면접원들로 구성된 국가실사진을 이용하여 이루어지고, SRC가 이를 조정·감독한다. 대략 115명의 면접원들과 6~12명 정도의 실사감독관들이 매년 데이터 수집에 종사한다.

13) 가구권으로 성립된 가구의 가구주는 심각한 장애자가 아닌 경우에는 남편으로 정의된다.

(1) 조사방식

조사방식은 계속적으로 변화하였다. 1968~1972년은 직접면접방식을 취했으나, 1973년부터는 비용절감을 위해 전화면접을 시행하였다. 1973년 이후에는 전화가 없거나(매년 약 500명) 특별한 이유로 전화면접이 불가능한 경우에는 직접면접을 수행하였다. 비용을 보다 줄이고 또한 전화로는 장시간 면접이 어려우므로 1973년부터는 면접시간을 줄였다. 직접 면접시에는 약 1시간 걸리던 것이 전화면접 실시 이후 평균 20~30분 소요되었다.

<표 2-13> 조사방법의 변화

연 도	조 사 방 법
1968~72	면접조사(기장조사)
1973~90	전화조사
1991~93	컴퓨터보조전화조사(CATI ¹⁴)로 점차 전환
1994~현재	컴퓨터보조전화조사(CATI)와 컴퓨터보조직접면접조사(CAPI ¹⁵)

2003년도 조사에서는 수기조사에서 컴퓨터보조면접조사(CAI)로 전환되었던 1991년 이래로 조사방법에 있어 가장 큰 변화를 찾을 수 있었다. PSID는 현재도 CAI를 사용하고 있지만 실제로 그 방법은 두개의 서로 다른 CAI 소프트웨어로 프로그램이 짜여진 다섯 개의 통합시스템이다. 아직도 대부분의 조사내용은 설문지 형식으로 수집되고 있고, 이렇게 조사된 내용들은 SRC 조사의 새로운 CAI 표준 소프트웨어인 'Blaise'라는 세 개의 시스템에 프로그램되었다. 그러나 응답자의 주거변화와 취업현황은 '경과이력목록(EHC)' 방법을 사용하는데, 이 방법은 조사자가 응답자의 답변에 내용은 EHC 목록에 바로바로 시간을 기록할 수 있고, 조금 더 편리하게 면접을 실시할 수 있게 되었다. 경과이력목록 조사방

14) CATI: 컴퓨터에 미리 입력된 조사대상자들의 전화번호를 이용해 컴퓨터가 임의로 전화를 걸고 전화통화를 통해 들은 응답결과를 면접원이 컴퓨터에 입력하는 방법

15) CAPI: 조사대상자가 컴퓨터를 이용해 직접 응답하는 방법

법은 조금 더 정확한 데이터 수집과 조사자와 응답자가 경직되지 않은 분위기에서 면접조사를 진행할 수 있게 하기위해 고안된 것이었다.

(2) 조사내용

PSID 데이터는 조사대상 가구의 가구와 개개인에 대한 수년간에 걸친 연구를 통해 방대하고 다양한 자료들을 가지게 되어 PSID 내용들을 나열하면 300페이지가 넘는다. 이중에는 매년 공통적으로 나타내는 내용들이 있는데 이를 PSID의 핵심조사항목들이라고 부른다. 이 핵심조사항목은 소득, 빈곤, 가족구성, 고용, 가사, 주거, 사회·경제적 배경 등이다. 그 밖에 연구과정에서 때때로 추가되는 항목들이 있는데 이를 추가항목이라고 부른다. 여기에는 공공요금, 교통이용 상황, 육아, 부(wealth), 보건, 그리고 친구, 친지사이의 시간 및 금전의 이전 등이 있다.

PSID 패널데이터 형성 초기의 조사항목에는 경제적·인구학적 분석에 초점을 두고, 크게 전체와 각 가구구성원의 소득원천과 소득액, 취업상태, 가구구성원 변화, 거주환경 등에 대한 실제적인 세부항목들이 포함되었다.

1985년 초부터는 아내의 출산과 결혼경력도 조사에 포함되었고, 이 밖에 식생활 지출, 가사노동 시간, 소비활동, 재산, 연금, 저축에 대해서도 조사하였고 1999년 이래로 건강상태에 대해서도 조사하기 시작했다.

비정기적으로 실시된 추가조사항목들을 살펴보면 패널조사 초기에는 표본가구의 주거상태, 주변 환경, 보육, 성취동기, 직업분류 경력, 그리고 퇴직 후 계획 등에 대해서도 조사하였다. 최근에는 교육, 군대경력, 건강, 친족관계, 그리고 가구의 재산 등이 특별조사항목에 포함되었다.

건강상태와 관련하여 추가된 조사항목은 국립노화연구원(National Institute on Aging)에서 55세 이상의 가구구성원을 대상으로 데이터 수집을 실시하였다. 세부내용은 일반적인 건강상태, 일상활동 상태, 재가복지(의료관련), 의료관련 지출에 관해 조사하였다. 1993년부터 1995년까지 PSID 연구는 노인에 대해 가구가 노부모의 건강을 돌보는데 필요한 지출과 그들을 위해 가족구성원들이 소비하는 시간에 초점을 두고 연구를 시작했다. 1984년, 1989년, 1994년, 그리고 1999년에 NIA의 승인을 받아 수집된 데이터들을 통해 PSID 패널가구의 재산과 저축에 관한 전망곡선을 산출할 수 있었다.

보육관련 조사는 1997년에 국립아동보건 및 인간발육연구소(National Institute of Child Health and Human Development)가 후원하고 PSID 표본가구의 0~12세 아동 약 3,600명을 대상으로 실시되었다. 이 조사항목에는 아동의 가정환경, 가족변천사, 가정과 학교에 대한 아이의 일지, 교육 및 주간보호 환경, 그리고 아이의 인지·감성·육체적인 기능에 대한 내용들이 포함되어 있었다.

〈표 2-14〉 PSID의 핵심조사항목

구 분	내 용		
소득원천 및 소득액	가족구성원들의 소득	사업소득 또는 농업소득	
	전문업 및 상업 소득	농업 및 조경업 소득	
	임대소득	주식배당	이자
	신탁 및 로열티	퇴직금	연금
	각종 사회보장제도로부터의 소득	위자료	
	직장으로부터의 비금전적 소득	육아지원금	
	친척 및 친지들로부터의 보조금	그 밖의 소득	
빈곤	가구 빈곤상황		
양식과 주거면에서의 공공지원	식권 사용실태 주거와 관련된 공공지원(공공주거사업대상자인지의 여부, 임대 의 경우 공공지원을 받는지의 여부, 난방비의 정부지원)		
기타 재정문제	연방세 납부 금액		
	자가 및 임대 이외의 거주에 대한 재정지원		
가족구성 및 인구학적 통계	결혼행위 및 혼인상태	임신	양자, 양녀
	자녀수(초 수 및 현재 같이 살고 있는 자녀수)		
	민족 인종		

〈표 2-14〉 계속

구 분	내 용		
고용에 관한 정보	주당 근로일수	실업기간	
	비경제활동일수	휴가(주 단위)	
	자신의 병 및 가족구성원에 대한 병구완으로 빠진 일		
	파업한 주의 수 등에 대한 월간 및 연간정보		
	주직장과 부직장에 대한 직종 및 산업		
	공무원 여부	임금률	주간근무시간 수
	주직장에 대한 노동조합 가입 여부		
	자영업의 경우 사업자 등록여부		
	근로경험: 총경험연수와 현직장에서의 경험연수		
	고용 상황: 면접시의 고용상태		
	구직활동 여부와 방법		
	이전의 고용상황 변화		
	고용주가 변동된 경우		
	고용주는 불변이나 직위가 변동된 경우		
직종 및 산업의 변동			
각 고용주하에서 일한 시작과 끝의 보수와 작업량			
고용주를 바꾼 이유			
가사종사 시간			
주거	주거하는 곳의 구조와 규모		
	소유상태(자가, 임대) 소유의 경우 집의 자산가치 남아있는 모기지(mortgage) 액수		
지역이동	지난해 동안의 이사(언제와 왜)		
	장래 이사계획(얼마나 확실한지 그리고 왜)		
	주거하는 곳의 주와 카운티		
	가구주가 성장한 곳 현재 주거하고 있는 곳 가구주가 살아왔던 모든 주들 가구주의 전직을 위한 이사 여부		
사회·경제적 배경	총교육과정	부모의 교육수준	자녀수
	인종 및 민족	아버지의 직업	
	부모의 빈곤상황		
보건, 종교, 군복무	가족구성원들의 건강상황 및 장애상태		
	종교	군복무 경험	
국가단위의 데이터	실업률	비숙련 노동자들의 임금	
	노동수요 상황		

자료: 신동균, 『미국 패널 데이터의 현황과 시사점』, pp. 22~23 재인용.

〈표 2-15〉 추가된 주제

연 도	추가된 주제 내용	
1968~72 1977~87	전기, 가스, 수도 등의 공공요금	
1969~86	통근 수단	
1968~72	거주 및 이웃 특성, 태도 및 행동양식, 직장외에 스스로 하는 활동, 저축, 가족구성원들의 장애상태, 임신 및 가족계획, 육아, 시간 활용	
1972	성취, 동기, 인식력(문장력 테스트)	
1973~74	육아	
1975	이웃에 대한 만족과 주거문제, 태도, 가구주의 장애상태	
1976	아내에 대한 면접 -고용경험 -임신과 가족계획 -직업의 특성(요구되는 직업훈련 포함)	-노동력 참가 실태 -육아 -태도
1977	육아, 가구주의 장애	
1978	직업훈련, 직업을 갖게 된 방법, 은퇴 계획 및 경험, 은퇴계획, 가구주의 장애상태	
1979	직장외에 스스로 하는 활동, 육아, 인플레이션의 영향, 저축, 은퇴계획, 가구주의 장애상태	
1980	위급할 때 시간적·금전적 도움, 식권·사회보장소득 수혜자격 여부, 인플레이션의 영향, 가구주의 장애상태, 확대가족, 저축	
1981~83	은퇴계획, 실업기간 및 노동력으로부터의 이탈기간, 연간 입원 상황, 가족구성원의 장애 및 병	
1984	부(여러 자산의 가치), 금전의 소득, 연금계획과 연금 수혜권, 은퇴계획, 유산, 저축, 직업훈련, 실업기간 및 노동력으로부터의 이탈기간, 가구주와 아내의 장애 및 병	
1985	아내에 대한 면접 -출산에 대한 회고 -양자, 양녀에 대한 회고 -대리부모 역할에 대한 회고 -결혼내역에 관한 회고 -전 교육과정에 대한 회고	-육아 -가사 -가족계획 -가구주, 아내의 장애와 병 -직업훈련
1986	모든 가족구성원들의 건강(일일활동 상황, 연간입원상황, 키와 몸무게, 흡연, 운동)	
1988	확대 가족들간의 연계(부모의 재정상태, 부모의 건강과 시간 및 금전적 지원)	
1989	부(여러 가지 자산의 가치), 저축행위(1984~89)	
1990	노령자들의 건강과 의료행위	

자료 : 신동균, 『미국 패널 데이터의 현황과 시사점』 pp. 23-24 재인용.

2003년도 설문조사 내용을 보면 이전 조사와 크게 달라진 점은 없다. PSID는 표본가구의 결혼과 출산, 가구 구성원들의 가구내 이동상황, 그리고 새로운 표본가구의 구성상황 등 인구학적 변화를 관찰하고 특별조사연도에는 동거인도 조사에 포함되었다. 또한 PSID는 주거상태, 주거비에 대해서도 조사하는데, 2003년도에는 지난 조사 이후의 소재지 변화에 대해서도 세부적으로 조사하였다. 또한 집에서의 컴퓨터 이용에 대해서도 추가적으로 수행되었다.

한편 2001년부터 노동인구의 노동시장 참가와 구체적인 취업현황 등 새로운 방법을 이용하여 수집되었다. 식료품비, 식권 사용 그리고 식품보급 안정도에 대해서도 조사하였다. 그리고 PSID의 주요 조사내용인 구성원 개개인과 가구전체의 구체적인 수입과 재산에 대해서 부모의 연금과 퇴직 후 계획을 토대로 상세하게 조사되었다.

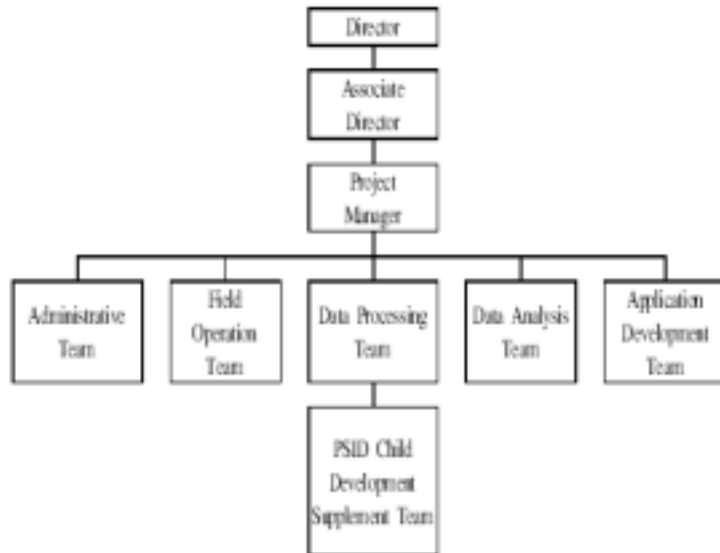
또한 지금까지 노부모의 건강과 만성질환 평가, 건강활동 그리고 전체가구구성원들의 의료비용에 대한 조사가 입원일수와 노부모의 일상생활능력(IADL)이 새롭게 추가되었다. 그러나 영양학적 지식에 관한 조사는 2003년에 제외되었다. 2001년부터 조사하였던 자선행사 참가에 대한 조사는 노부모의 자발적인 참여에 관해서도 조사함으로써 그 범위가 좀더 확대되었다. PSID는 지난 2년 동안 사회복지프로그램의 참가에 대한 조사를 계속해 왔고, 2년 주기의 조사에서 지나쳤던 소득과 노인들을 위한 노력의 총체적인 사항들을 계속 조사했다.

라) PSID 관리체계

(1) PSID 조직

PSID는 체계적인 조직 아래에서 운영되며, PSID 운영을 위해 행정팀, 조사현장 운영팀, 자료 처리팀, 자료 분석팀 등이 구성되어 운영하고 있다. PSID 패널조사가 진행되는 과정에 따라 각 팀들이 활동하고 Project Manager가 이를 총괄한다.

〈표 2-16〉 PSID 조직도



(2) 조사 및 관리

면접원들과 실사감독관이 데이터 수집에 참여한다. 실사감독관들은 면접원들을 선발하여 교육시키고 면접과정을 감독한다. 또한 면접의 질과 완결성, 정확성 등을 검토하면서 면접원과 응답자 사이에서 발생하는 제반문제들을 해결한다. SRC 사무국은 감독관들과 면접원들의 실사수행에 대한 모의실험을 통하여 실제의 면접이 원활하게 수행되도록 하며, 실사상에서 감독관들 및 면접원들로부터 질문에 답하며, 면접자료들의 배포 및 수거를 맡는다. 나아가 매일 감독관들 및 면접원들로부터 접수되는 설문지와 표지들을 기록, 관리하기 위한 표본통제관리 시스템을 유지한다. PSID 참모들은 면접원들의 선발 및 교육을 도와주고, 응답자들의 선정 및 면접에 관한 실사 사무국으로부터의 질문에 답하며, 모의전화면접을 수행하는데 도움을 준다. 또한 면접원 교육과정을 참관하고 필요시 감독관 및 면접원들과 의견을 교환하며, 실사사무국으로부터의 표본통제관리 시스템에 대한 보고를 검토한다.

새로 선발되는 면접원의 수는 매년 다르다. 추가적인 면접원의 고용은

표본가구들이 위치해 있는 지역에서 이루어진다. 사무국은 새로운 면접원들을 선발하고 실사 감독관에 의한 기본면접 교육을 마련해 준다. 기존의 면접원들을 포함하여 모든 면접원들은 실사전에 사무국의 사전회의에 참석한다.

1968년의 PSID 표본은 200개를 상회하는 1차 표본지역(primary sampling area)들에 군집(cluster)되어 있었다. 이는 데이터 수집의 용이함과 이에 따른 비용절감을 위한 것이었다. 그 후 시간의 경과에 따른 PSID ‘표본구성원’들의 이주로 표본은 미국전역 50개주 모두와 여러 외국들도 포함하게 되었다. 연평균 250명의 새로운 가구주들을 면접하였다. 전체 PSID 가구들 중에서 약 92%의 가구들을 전화로 면접하였다. 나머지 8%는 전화가 없거나 청력장애 및 건강상의 이유 혹은 개인적인 사유로 직접 면접을 실시하였다. PSID 초기에는 면접원들을 지리적으로 인접한 응답자들에게 할당하였다. 직접면접은 지금도 이 원칙에 의해 수행된다. 그러나 1985년 장거리 전화요금을 절약하기 위해 전화면접원들도 재배치하였다. 미국의 장거리 전화요금 구조로 인하여 이러한 재배치는 주(state)내에서의 장거리 전화면접에 대한 각 주들 사이의 장거리 전화면접의 비율을 높였으며, 장거리 전화 응답자들은 그 이전의 면접원들로부터 분리시키게 되었다. 하지만 어느 시기를 막론하고 가능한 응답자들에게는 고정적인 면접원들을 접촉시키는 방향으로 노력했다.

1991~1993년까지 3년간 PSID는 컴퓨터보조면접 방식으로 면접방식을 전환하고 있었다. 첫 단계는 SRC 전화면접진에 의하여 가능한 많은 면접들을 중앙화(centralize)하는 것이었다. 점차 종이와 연필에 의한 면접으로부터 컴퓨터보조전화면접(CATI)과 컴퓨터보조 직접면접(CAPI)으로 바뀌고 있다.

(3) 감독

면접원들이 면접을 완료한 설문지들의 일정부분을 감독관들에 의해 면접이 실제로 이루어졌는지가 검토된다. 즉, 감독관들은 일부 설문지들을 선택하여 해당 응답자들을 다시 접촉한다. 숙련된 면접원들에 의해 완성된 설문지들의 경우 약 5%정도가 재접촉되며, 새로운 면접원들에 의해 수행된 설문지들의 경우에는 약 10%정도가 재접촉된다. 감독관들은 이

재접촉을 통해 응답자들에게 약 10개의 질문을 하는데 여기에는 면접이 얼마나 잘 진행되었는지와 면접에 소요된 시간 및 설문지에서 추가한 질문들이 포함되어 있다.

(4) 응답률 극대화를 위한 노력

표본의 횡·종단면적 대표성은 패널데이터의 생명과도 같다. 이에 대해 PSID의 대표성 확보를 위한 노력들을 간략하게 살펴보았다.

미국 노동시장의 가장 특징적인 면 가운데 하나는 직장간의 노동이동이다. 따라서 최초의 표본가구들을 계속해서 추적조사하기 위해서는 상당한 노력이 필요하다. 이는 결국 응답자의 협조를 얻어 내기 위해 어떤 종류의 노력을 어느 정도 기울이는가의 문제이다. PSID의 응답률을 높이기 위해 행한 노력은 아래와 같다.

- 이사진 표본가구들을 추적하고 응답을 거부하는 가구들을 설득하기 위해서는 상당한 시간이 필요하다. PSID의 실사기간은 3~9월로 상당히 길어서 최대한의 추적을 가능하게 한다.
- 각 응답자들에게 \$55을 지급하는데 이는 면접에 응해준 것에 대한 보상과 주소가 변경된 경우 PSID가 제공한 주소변경카드를 반송하는데 드는 경비가 포함된 것이다.
- 매년 분석된 PSID 데이터의 내용들과 이들이 정책토론에 어떻게 사용되었는지를 요약한 책자를 모든 응답자들에게 우송한다.
- 응답을 거부하는 사람들에게는 설득하는 내용의 개별편지를 보낸다.
- 모든 면접의 마지막에는 응답자들이 이사할 경우 그들의 새로운 연락처를 알 수 있는 친구나 친지들의 이름과 전화번호를 받아낸다. 이러한 정보를 보유하고 있는 참모진들은 면접원들의 추적상의 애로사항이 있을 때 적극 협조한다.
- 전화면접이 어려운 경우에는 직접 면접을 실시한다.
- 정기적인 특별 우편접촉(수표, 응답자들의 보고, 다음 면접에 대한 공지사항 등)으로 응답자들이 이사하는 경우 변경된 주소를 빨리 포착할 수 있게 한다.

2) 영국 : The British Household Panel Survey(BHPS)

가) BHPS 개요

1991년 9월에 시작한 영국의 가구패널조사(The British Household Panel Survey: BHRS)는 사회·경제구조의 변화 추세를 관찰하고 이에 대한 정보를 수집 및 분석하여 21세기 영국의 정치, 사회, 경제 구조를 예측함으로써 각종 정책 수립의 기초 자료로 활용될 수 있도록 고안되었다. 이 조사는 매년 실시되면, 2004년 9월 현재 14차년도 조사(wave 14)가 시작되었다.

BHPS의 주요 목적은 영국의 개인 및 가구 단위에서 일어나는 사회적·경제적 변화에 대한 이해를 증진시키고, 사회·경제적 변수 범위와 관련한 원인 및 결과와 변화를 확인하고 유형화하고 예측하기 위한 것이다.

BHPS는 Essex 대학에 있는 Institute for Social and Economic Reserch(ISER) 내의 SERC Research Centre oin Micro-social Change와 UK Longitudinal Studies Centre에서 주관한다.¹⁶⁾¹⁷⁾ ISER는 조사의 기술적 측면 즉, 조사 설계, 질문지 발전, data 유효화 및 가중치 등에 대한 책임을 지고 있다.

그리고 실사는 민간조사기관인 NOP(National Opinion Poll) Market Research Ltd. 가 1991년부터 수행하고 있다.¹⁸⁾ BHPS의 실사는 두 단계로 수행되고 있다. 4월에서 5월까지 1,000명을 표본으로 한 예비조사(pilot survey)가 실시되고, 여기에서 질문 내용들이 검토되었다. 본 조사는 9월에서 12월까지의 4개월에 걸쳐 실시되었으며, 응답가구는 £10에 해당하는 선물증서(gift voucher)를 받는다.

16) 1999년에 만들어진 ULSC(UK Longitudinal Studies Centre)는 1989년 ESRC의 보조금으로 만들어진 ESRC Research Centre on Micro-Social Change의 연구 자원 요소의 연속선에 있는 것이다.

17) Institute for Social and Economic Research(ISER) 내에는 ④ ESRC Research Centre on Micro-social Change, ⑤ UK Longitudinal Studies Centre, ⑥ European Centre for Analysis in the Social Sciences(ECASS) 등의 Centre들이 있다.

18) 북아일랜드에서의 현장 조사는 NISRS에 의해 수행되고 있다.

BHPS의 주된 활동들은 전체적으로 UK Economic and Social Research Council에 의해 재정 지원을 받으며, 북아일랜드의 조사는 Eurostat에 의해, 11~15세 조사는 The UK Health Education Authority에 의해 지원 받는다.

BHPS에는 소득, 소비, 노동시장활동, 건강, 훈련, 연금, 주거상태 등 경제 활동 전반에 걸친 다양한 질문들 뿐 아니라 시간의 흐름에 따라 개인의 삶이 어떻게 변화하는지를 파악할 수 있는 내용들로 구성되어 있다. BHPS는 광범위한 사회 과학 분야를 위한 그리고 많은 분야에서 학제 간 연구를 지지하기 위한 연구 자원으로써 설계되었다.

BHPS는 약 5,000가구 내에서 10,000 명 이상 면접하여 전국적으로 대표성을 띤 표본으로 16세 이상 성인 각각에 대해 매년 조사를 실시한다. 1991년 1차년도 조사에 면접을 성공한 가구 수는 5,130 가구로 가구 구성원 중 16세 이상의 모든 사람이 조사 되었다.

나) 표 본

BHPS는 5,000 가구 이상에서 전체 대략 10,000 명의 개인을 면접하는 전국적으로 대표성을 띤 표본으로 16세 이상 성인에 대해 매년 조사를 실시하도록 설계되어졌다. 그리고 같은 개인들이 원가구에서 분리되지 않는다면 다음 년차 조사에서 다시 면접되어지며, 가구에 새로 들어온 모든 성인 또한 면접되어 진다.

BHPS 1차년도 조사의 첫 표본은 우편번호 주소록 파일(Postcode Address File ; PAF)로부터 추출한 8,167개로 구성되었으며, 면접은 그 주소에서 발견된 모든 개별 가구에게 시도되었고, 응답 가정 안에서 열거된 모든 개인은 종단 표본의 일부가 되었다. 이 모든 표본 숫자가 OSMs(Original Sample Members, 원표본구성원)이다. 연속된 조사의 표본은 1차년도 조사에 면접했던 개인이든 아니든 관계없이 가구에 살았던 사람을 포함한 모든 가구에 있는 성인으로 구성된다.

1차년도와 그 이후 조사에서 표본 적격성은 잉글랜드, 웨일즈, 카렐라니아 운하 남쪽 스코틀랜드 안에 거주하는 것에 따른다. 1차년도 이후 표본에 들어가는 새로운 자격 조건은 ① OSMs(원표본구성원)에서 태어난 아기, ② OSMs(원표본구성원)이 한명 이상의 새로운 사람이 있는

가구로 이동했을 때, ③ 한명 이상의 새로운 사람이 OSMs(원표본구성원)으로 들어 왔을 때 등이다. 즉, 표본은 출생, 혼인, 동거 등으로 인한 새로운 가구원으로 늘어나기도 하고, 반대로 사망, 거부, 조사 지역 밖으로의 이사 등으로 인해 줄어들기도 한다.

BHPS은 가구 형성 및 해체를 추적하는 가구 패널조사인데, 그 안의 개인은 표본을 떠나기도 혹은 들어오기도 한다. 조사 시 개인이 면접 자격이 있는가 하는 것은 다음의 3가지 표본 구성원 지위(OSMs, TSMs, PSMs)와 관련이 있다.

첫째, 원표본구성원(OSMs; Original Sample Members)은 1차년도 조사 가구의 모든 구성원을 이르며 더불어 조사 이후에 출생한 그들의 친자를 포함한다. 원표본구성원은 항상 면접 가능한 자격이 있다. 둘째, 일시적 표본구성원(TSMs; Temporary Sample Members)은 연구를 시작한 이후에 원표본구성원(OSMs)과 가구를 형성한 사람을 이른다. 이들은 원표본구성원과 함께 사는 동안만 면접 대상 자격이 있고, 그들이 떠난다면 면접 대상의 자격이 주어지지 않는다. 셋째, 영구적 표본구성원(PSMs: Permanent Sample Members)으로 일시적 표본구성원이 원표본구성원과 충분히 강하고 지속적인 유대를 갖는다고 간주된다면 PSMs이 될 수 있다.

그리고 BHPS에는 1차년도 조사 이후에 추가로 만들어진 3가지 표본 ① ECHP(유럽공동체가구패널) 하위 표본, ② 스코틀랜드 및 웨일즈 확대 표본 그리고 ③ 북아일랜드 확장 표본 등이 있다.

첫째, 7차년도 조사부터 BHPS는 UK의 『ECHP』(유럽공동체가구패널)에 데이터를 제공하기 시작했다. ECHP 조사는 EU 안에서 수입, 삶의 조건, 주거, 건강, 노동 등에 대한 비교가능한 미시 수준(개인/가구)의 데이터를 제공해 주는데, BHPS은 ECHP를 도입하였고, 이에 필요한 몇 가지 질문을 변경 및 첨가하였다.

둘째, 스코틀랜드와 웨일즈 확대 표본이다. 스코틀랜드 확장 표본은 9차년도 조사(1999) 때 ESRC에 의해 재정을 지원 받아 추가되었다. 스코틀랜드 북쪽 지역을 포함하여 1,500 가구로 확대 되었고, 이렇게 확장된 표본에 대해 국가 정체성이나 정부 구조에 대한 태도 등의 새로운 질문들이 추가되었다.

셋째, 북아일랜드 가구 패널조사(NIHPS)로 북아일랜드에 있는 새로운 표본들이 11차년도 조사(2001)부터 첨가되었다. 이 표본 조사를 위해 ESRC와 북아일랜드 정부 부서에서 함께 기금을 마련하였다. 그리고 북아일랜드에서의 현장조사는 NISRA에 의해 수행되었다.

BHPS에서 표본 탈락(attrition)은 각 조사연차 사이에 가구 차원에서 발생하거나 개인 차원에서 발생한다. 그리고 항목 무응답(item nonresponse)으로도 발생할 수 있다. 이러한 패널 탈락을 최소화하여 지속적으로 표본의 대표성을 보장하기 위한 방법은 다음과 같다.

- 구성원들이 영국 내 어디로 이사 가든 따라 다닌다.
- 원패널구성원(OSMs)에 의해 형성된 가구의 새로운 구성원을 패널 안에 포함시킨다.
- 효과적인 현장 조사를 실시한다.
- 면접자들에게 비디오 시청각 훈련을 실시한다.
- 특별 보고서 및 편지 등을 통하여 패널 구성원들과 정기적으로 연락한다.

대부분의 표본 조사에서는 표본의 추출 확률이 일정하지 않기 때문에 이러한 상이한 표본추출확률을 가중치¹⁹⁾를 통해 조정하여야 한다. 표본 조사 자료의 결과가 대표성을 가지려면 적절한 가중치를 부여하는 것이 필요하다. 패널조사에서는 초기 조사(wave)에서의 표본추출확률이 동일하다고 하더라도 조사(wave)가 진행됨에 따라 표본 탈락이 발생하기 때문에 가중치를 조정해 나갈 필요가 있다.

BHPS는 영국에 있는 개별가구의 종단조사로 1차년도 조사에서 목표 표본을 5,000가구와 10,000명으로 설정하였다.²⁰⁾ 패널조사에 포함되기 위한 가구의 초기 선별은 2단계집락확률설계(two-stage clustered probability design)와 계통(체계)표집(systematic sampling)을 사용하였다. 표본 단위(sample unit) 선정을 위한 틀은 Caledonian 운하(wmr, 북아일랜드

19) 가중치(weight)는 어떤 대상이 속해 있는 집안에서, 다른 대상에 비해 상대적인 중요성을 수치로 나타내는 값을 말한다(자료: 통계청).

20) 이는 미국 PSID와 독일 GSOEP에서의 1차년도 조사 표본수를 고려하고, 또한 조사의 성격상 다목적 조사이기 때문에 표본 내에 포함되는 다양한 하위집단의 표본수가 일정 규모 이상이 되어야 한다는 점을 고려한 결과이다.(P. 15 인용).

드는 제외)의 남쪽 영국에 해당하는 PAF(Postcode Address File, 우편번호주소록)이다. 이는 대규모 정부조사에서 일반적으로 사용되는 틀로 효율과 비용의 균형이 고려되었다.

다) 조 사

BHPS 질문지는 사회경제구조의 변화를 연구하는데 적합하도록 구성되어있는데 주요주제는 가구동학(動學), 노동시장, 소득과 부의 분배, 주거, 건강, 그리고 사회경제적 가치 등이다. 구체적인 내용은 아래의 표와 같다.

〈표 2-17〉 BHPS의 주요주제

구 분	내 용
가구動學	가구구조의 변화: 가구형성과 해체의 원인, 결과 가구경제 의사결정 가구소비 양식
노동시장	직업이동, 노동시장 구조
소득과 부의 분배	생활수준의 변화 추이 생활주기 효과 소득과 가족구성 부의 축적과 저축
주거	주거비용 주거이동의 원인 및 결과
건강	건강과 가구경제와의 관계 의료서비스의 이용실태와 의료서비스 이용상의 불평등 정도 가구원의 건강상태와 질병의 종류
사회경제적 가치	사회경제적 지위와 가치(행위)와의 관계

BHPS 질문지는 가구용 질문지, 개인용 질문지, 자기기입 질문지, 대리설문자가 따로 마련되어 있는 것이 특징이다. 가구용 질문지는 가구원 중에서 주거상황에 대해 가장 많이 알고 있는 가구원이 응답하며, 개인용 질문지는 만16세 이상의 모든 가구원이 응답한다. 자기기입 질문지는

면접조사를 마친 모든 가구 구성원이 기입하며, 대리질문지는 선원 등의 이유로 집을 떠나있는 가구원이나 고령으로 응답이 불가능한 경우 다른 가구원이 대신 하여 응답하는 질문지이다. 대리질문지는 응답을 거부한 경우에는 사용할 수 없고 이를 사용하는 경우는 최소한 4번 이상의 접촉시도가 있어야 허용된다.

BHPS은 1991년부터 1998년까지는 서식 질문지(paper and pencil questionnaire)를 이용하여 면접자가 면접단계식 조사로 자료를 수집하였는데, 이때는 면접자 관리 질문지와 짧은 자기완성 질문지 두 가지 양식을 포함하였다. 1994년부터 오디오 테이프를 이용한 자기완성 질문지가 11~15세 아동에게 도입되었다. BHPS는 9차년도 조사 이전에는 서식(PAPI)이었으나, 9차년도 조사에서 컴퓨터 지원 개인면접(CAPI; Computer Assisted Personal Interviewing)으로 전환하기 시작되었다. 이 시스템에서는 ‘모른다’와 ‘응답 거부’ 항목이 있기는 하지만 응답을 해야만 다음 질문으로 넘어 가게 되어 있다.

(1) 질문항

BHPS 질문지의 질문항은 핵심 요소, 순환 요소, 가변 요소로 구성되어 있다. 첫째, 1차년도 조사에서 질문된 많은 질문들은 다음 년차 조사에서 되풀이 되어지는데 전체 조사에서 반복되는 것을 ‘핵심’(core) 질문이라고 한다. 이것은 조사의 심장부이며, 조사 목적인 순수한 변화를 연구하도록 도와주는 부분이다.

- 주거, 소비, 이웃 특성
- 가구 형성, 가사 노동
- 교육 및 훈련
- 노동 시장 행동, 현재 직업, 일년간의 직업력
- 건강, 활동 제한, 건강 서비스의 이용
- 사회·정치적 가치, 사회 참여 및 네트워크
- 수입: 최근 및 일년 추정치, 비금전적 척도

둘째, 일부 변수는 격년 혹은 순환 주기(2~3년에 한 번)에 따라 조사되는 주제인데 이를 ‘순환rotation’ 질문지라고 한다. 시간에 걸친 큰 변화를 기대

하지 못하기 때문에 매년 조사할 필요가 없는 주제를 포함하는데, 도덕성 및 종교에 대한 태도, 삶의 만족도, 정부의 역할 및 책임, 환경 등이다.

셋째, 일부 질문군은 패널조사 과정에서 딱 한번 물어보는 ‘비핵심 혹은 가변 구성요소’(non-core or variable components)이다. 통상 한번만 질문되어지는 가령 ‘당신은 언제 학교를 떠났습니까?’라든지 ‘당신은 어디에서 태어났습니까?’와 같은 질문이다.

(2) 질문지 패키지²¹⁾

① Cover Sheet

면접자는 cover sheet에 면접대상자에게 통화한 기록 및 횟수, 대상자의 주거상태 및 유형에 대한 면접자 자신의 관찰 내용, 가구(시설) 특징이나 면접 거부 이유, 전체 및 16세 이상 가구원의 수, 가구 위치도 등 최종가구결과(final household outcome) 등을 기록한다. 그리고 Household Grid를 작성하고, 가구 구성원의 변화 즉, 이사 및 전입 등과 그 이유와 날짜 및 동거 기간, 시설 입소 등 가구 변화를 기록한다.

② 가구 질문지(Household questionnaire)

가구 질문지(가구 구성표(A household composition form))는 한명 이상 면접을 성공한 모든 가구에서 실시되며, 대부분 가구 성인구성원과 면접자의 첫 번째 만남에서 이루어진다.

가구 질문지의 내용은 주거(accommodation)에 대한 질문이 주를 이루는데 주거지 크기, 주거 상태 및 문제, 소유 상태, 보유 및 임대 기간, 주거비용 방법(현금, 모기지), 이전 소유권, 주거 급여(임대 리베이트, 임대수당) 수혜 여부 등을 포함하며, 텔레비전·비디오 레코더·식기 세척기 등의 가전기기 소유 여부와 자동차 소유 여부 및 주용도 등을 묻는다.

③ 개인 질문지(Individual questionnaire)

가구의 모든 16세 이상 어른에게 수행되는 개인 질문지는 약 35~45분이 소요된다.²²⁾ 개인 질문지가 포괄하는 내용은²³⁾ 이웃 관계, 개인 신상, 주거

21) 일차년도 질문지 참고

22) 개인 면접이 끝나면 면접자는 면접시 다른 사람이 함께 있었는지, 그로 인해 어떤 영향을 받았는지, 응답자가 어느 정도 잘 협력하는지, 이후 추적 조사에 응하려 하는지, 면접시 장애 요인(시력 및 청력 변화)이 있는지 그리고 면접시 특이점 및 주의점 등 자신이 관찰한 내용을 기록한다.

이동, 건강과 보호(caring), 최근 고용과 수입, 지난해 고용 변화, 출산, 혼인 및 관계력(W2에서만), 고용 지위 변화(W2에서만), 가치와 의견, 가계 재정과 조직 등이다. 개인용 설문지의 구체적인 내용은 <표 2-18>과 같다.

④ 자기완성 질문지(Self-completion questionnaire)

자기완성 질문지는 개인 질문지에 응답한 모든 반응자를 대상으로 하며 약 5분이 소요된다. 응답하는 동안 타인이 있으면 특별히 영향 받기 쉬운 주관적이거나 태도와 관련된 질문이나 부가적인 비밀보장이 요구되는 민감한 질문들이 포함된다. 그래서 자기완성 질문지는 특히 비밀과 익명성이 보장된다.

핵심 질문으로 스트레스, 걱정, 억압, 역량, 강점, 능력, 행복과 불행 등에 대한 주관적 만족이 포함된다. 그리고 순환 질문에는 태도 측면(가족에 대한, 남성/여성 역할에 대한, 종교에 대한), 사회적 지지망, 삶의 만족도 등이 있다.

⑤ 대리 질문지(Proxy questionnaire)

대리 질문지는 현장조사 기간 동안 가구 구성원이 부재하거나 혹은 면접을 완수하기에 너무 늑거나 약한 경우 그 사람의 정보를 모으기 위해서, 배우자 또는 성인 아동 등 다른 가구 구성원을 통해 작성하는 것이다.²⁴⁾ 그리고 「전화 질문지」는 대리 질문지로부터 발전된 것으로 기관에 의해 고용된 경력 면접자가 직접 면접을 시도하려는 모든 노력이 실패되었을 때 사용하는 것이다.

⑥ 아동 질문지(Youth questionnaire)

아동 질문지는 조사 대상 가구 내에 포함되는 11세에서 15세 사이의 아동에게 그들의 건강, 가족, 희망과 관심 등에 대해 조사하되 직접 기입하는 방식으로 진행된다.

23) 개인 면접의 결과는 ㉠ 완수(Full interview), ㉡ 대리(Proxy), ㉢ 거부(Refusal), ㉣ 부재하고 대리 질문도 못함(Absent, no proxy), ㉤ 접촉 실패하고 대리 질문도 못함(Other non-contact, no proxy), ㉥ 있었으나 면접 불가하고 대리 질문도 못함(Present but unable to be interviewed, no proxy) ㉦ 16세 이하 어린이(Child under 16) ㉧ 기타로 구분된다.

24) 면접자는 다음 년차 조사를 위해서라도 개인이 면접을 거절하거나 접촉이 되지 않은 원인에 대해 구체적으로 기록하고, proxy 면접조차 부적절했거나 불가능했던 이유에 대해서도 기록한다.

〈표 2-18〉 개인용 설문지의 내용

	영역	내용
핵심	이웃관계와 개인의 인구학적 정보	출생지, 집과 이웃에 대한 주거 만족도, 이사 이유, 인종, 교육 배경과 학력, 최근 교육·훈련, 지지하는 당, 혼인 상태에 있어서 변화
	현재 고용	고용 지위(실업 및 구직, 자영업, 민간 부문·공공 부문 등), 사업 특성/업무, 근무지, 직장 규모, 직장까지 걸리는 시간과 교통수단, 근무 연수, 근무 시간/ 초과근무시간, 노동조합 소속 여부, 전망과 훈련 및 야망, 퇴직금/연금 체계, 일에 대한 태도, 인센티브, 급여, 공제, 보육 시설, 구직 활동, 경력 기회, 보너스, 급여 관련 활동
	재정	급여, 수당, 연금, 집세, 저축, 이자, 배당금 소득, 연금 계획, 조축 및 투자, 물질적 풍요, 소비자 신뢰, 내적 이전, 외적 이전, 개인 지출, 배우자 역할, 가구 지출, 육아, 청구서, 일상 소비, 자동차 소유 및 사용, 차의 가치, 면접 특성, 기타
	건강과 보살핌	개인 건강 상태, 고용 제약, 의사 방문 빈도, 병원 이용, 건강 보험·복지 서비스 이용, 사회복지 서비스, 전문의 진단, 건강 검진, 흡연, 친척이나 가족의 보살핌, 타인 보살핌에 드는 시간, 사적인 건강 보험, 일상생활 활동 가능성
순환	건강과 보살핌	건강 보호를 위한 비용/지불에 대한 태도
	가치와 의견	부의 분배, 사회 정의, 정부의 역할과 책임, 환경, 가구 지출의 관리, 지도자, 종교, 노동 조합, 사회적 계층, 사회적·이해 집단 활동에의 참여 여부, 레저 활동, 국가 정체성, 영국 정부의 자치와 지방 권한
가변	일생 혼인 지위 변화(w2)	혼인 횟수, 혼인 날짜, 이혼/과부/ 별거 날짜, 혼전 동거
	일생 임신과 입양 경험(w2)& w8	태생일, 사망일, 입양일
	일생 동거 경력 (w2)&w8	시작과 끝날, 파트너 수
	가치와 의견	자녀에 대한 열망, 중대한 사건, 삶의 질
	일생 고용의 변화(w3)	직장 첫날과 끝날, 노동 지위, 분야, 직무의 특성
	부와 부채(w5, w10)	투자와 저축, 채무
	건강과 보살핌	아동 건강, 다른 건강척도 : SF36(w9)
	이웃 특성(w8)	지역사회와 이웃, 지방 복지서비스의 질
	범죄(w7, w12)	지역에서의 범법적 활동, 범죄 인지 정도
	컴퓨터와 전산 처리(w6,7,12)	소유 및 이용
	고용(w9)	국가최저임금, 노동 억제 요인, 노동 정향
	노화와 은퇴(w11)	은퇴 결정, 삶의 질, 가족 지지망
	평생 고용지위 변화(w2)	시작과 끝나는 날짜, 고용 지위
자녀와 양육(w12)	학교 선택, 교육적 열망, 양육 스타일	

라) 관리체계

Centre는 현장 조사(Field Work)를 수행하도록 의뢰된 「NOP 사회·정치부」(MAI의 일부)와 함께 자료 수집의 모든 측면을 밀접하게 같이 작업하고, 효과적인 데이터 질과 적절한 반응을 보장하기 위해 기획되고 동의된 일련의 조사절차를 수행한다.

NOP는 현장 조사, 에디팅, 코딩, data-entry 등과 동시에 모든 조사 도구의 설계에 대해 조언을 한다. 작업 설계, 면접자 지표의 산물과 부가적인 브리핑 내용들의 설계 및 산물 등에 대한 일차적인 책임은 Centre에 있지만 NOP는 스태프들에게 에디팅과 코딩 결정에 대해 조언해 주어야 한다. Centre는 현장 조사 실무의 특수화, 에디팅과 코딩을 통하여, 질(quality) 통제에 주요한 역할을 한다.

(1) 면접자

1차년도 조사에서, 표본의 250 지역을 담당하기 위해 243명이 면접자가 고용되었다(대체로 1인이 1영역 담당). BHPS에 요구되는 성격 때문에 평균 이상 수준의 경험과 능력의 면접자가 고용되었고, 이들 상당수는 수년간 BHPS 조사에 관여하였다. NOP와 지방 NOP 현장 담당자에 의해 주어진 이사자의 구체적 정보를 통해 표본 영영을 유지하고자 한다. 금년 조사와 후년 조사 사이에 개인을 추적하는 일에 많은 노력을 한다. 면접자는 조사 과정에서 의문이 있을 때 언제든지 지역 지도감독관(area supervisor)에게 연락하여 해결하거나 그것이 여의치 않을 때는 NOP나 Essex Institute에서 도움을 받는다.

(2) 현장 조사 과정

1차년도 조사는 9월 3일에 시작하여 12월 중순에 끝났다(1992년 초에 일부 약간의 추후조사). 표본과 접촉하기 전에, 면접자는 모든 표본 주소로 Centre에 이름으로 「소개 편지」를 발송하는데 이때 조사의 목적을 설명한 간단한 ‘리플릿’을 함께 동봉한다. 면접자는 소개 편지를 발송한 후 일주일 안에 연락을 한다. 모든 참가구는 나중에 보다 자세한 ‘브로슈어(안내서)’를 받는데 이는 조사에 대한 자세한 정보를 담고

있으며 더불어 조사 참여자에 대한 감사를 담고 있다. 반응자에게 보내어 지는 서류의 복사본은 ESRC 조사 센터에 남겨야 한다.

면접자는 ‘접촉 불가’(non-contact) 처리를 하기 전에 개별 표본 주소에 최소한 6번의 전화를 시도해야 한다. 면접자는 가능하면 더 많이 전화를 하도록 하고, 만약 Centre가 참여를 거부한 가구에 대해 변경하고자 한다면, Centre 이름으로 「특별 변경 편지」를 보낸다.

이후 연속적인 조사에서 일반적으로 현장 조사 기간은 9월 초에 시작되고 대부분은 12월 초에는 마치는데 이사자 추적, 거부자 변경, 면접자수 부족 등으로 인해 조사 과정이 지체가 되어 다음 해의 첫 3달까지 수행되기도 한다.

완강하게 거부하는 사람을 제외하고 모든 자격 있는 응답자들에게 면접자에 의한 접촉이 이루어지기 전에 「advance letter(사전 편지)」²⁵⁾가 보내어 졌다.

면접자들은 그들 표본 가구의 모든 자격이 되는 구성원에게 「소개 편지(introductory letter)」를 제공해야 하고 가구에서 요구하기 전에 우편 발송을 해야 할 책임이 있다. 「advance letter」는 16세 이상 조사 대상 자격이 되는 개인에게 조사 전에 보내지는데 이때 선물 증서(gift voucher)가 함께 주어진다(£10, 단 아동 면접에서는 £5). 또한 여기에 ‘주소 변경 카드(change of address card)’와 ‘팸플릿’을 동봉한다. 이때 편지와 리플릿에는 ‘무료전화 번호’를 기재하여 더 많은 정보를 얻거나 응답자가 면접시간을 바꿀 수 있도록 기회를 제공한다. 팸플릿은 주로 조사에 새로 참여하는 사람에게 조사를 설명하기 위한 목적이 있다. 그리고 면접자들은 advance letter를 보낸 뒤 기억이 생생한 며칠 이내로 전화를 하되, 여러 가지 질문에 대비하여 조사 관련 내용을 명확히 숙지하고 있어야 한다. 그뿐 아니라 면접자들은 여분의 선물 증서(gift voucher)를 가지고 있다가 조사에 새로 들어오는 사람이나 이전 조사 후에 받지 못했다는 사람에게 준다.

조사의 질을 보장하기 위하여, 모든 면접자들은 면접 초반에 그들 지도감독관(supervisor)을 따라 다니며 그들이 하는 면접처럼 완수해야 한다.

25) advance letter는 lead letter라고도 불리며 면접자가 조사 대상자에게 연락하기 전에 보내는 편지를 말한다.

지도감독관은 2주간의 현장 과정 모니터를 통해 면접자를 평가한다. 두 가지 방식으로 면접의 질을 통제하는데 면접 전체를 완수한 가구 가운데 약 10%에 우편 연락을 하여 소환하거나 혹은 가구에서 일어난 면접과 수행된 개인 면접을 점검하기 위하여 지도감독관이 전화로 직접 확인하는 방법이다. 그리고 Centre는 전화로 질 통제 과정을 보완한다. 이러한 점검은 데이터의 질을 통제할 뿐 아니라, 면접 시간 및 내용, 개인적이고 민감한 정보에 대한 걱정, 기타 전반적인 인상 등 BHPS 면접에 대한 응답자의 반응을 알아보는 방법으로도 활용될 수 있다.

(3) 거부 변경(refusal conversion)

각 조사에서 Centre와 NOP는 거부 또는 다른 형태의 무응답(non-response)으로 인한 탈락(attrition)을 최소화하기 위해 철저한 거부 변경(conversion) 과정을 수행한다.

3차년도 조사에서 새로운 거부자가 나타났을 때, 다음과 같은 일련의 절차가 시행되어졌다.

- 면접자는 cover sheet 안에 있는 「거부 양식」을 이용하여 거부 특성 및 이유에 대해 구체적으로 질문한다. 그리고 각각의 지속적인 변경(conversion) 시도를 이 양식 혹은 연속 sheet에 기록한다.
- 경력 많은 면접자들이 참여 변경(conversion)을 위한 설득을 한다.
- 변경(conversion)이 성공하면 면접자에게 보너스를 준다.
- 일부 상황에서 면접자는 응답자가 변경(conversion)하기 전에 바우처(상품권)를 발행한다.

면접자들은 응답자의 거부 형태별로 다른 대처 반응을 해야 한다.

- 거부의 뚜렷한 이유가 없는 경우(‘귀찮다’, ‘너무 바빠다’, ‘지루하다’, ‘지난번에 바우처를 받지 못했다’ 등) - 이런 반응자들에게는 <변경 편지>를 보내거나, 면접 완수시 주는 바우처를 다시 한번 발행한다.
- 거부예의 강한 이유가 있으나 아직 결정하지 않은 경우(‘아픈 부모님 간호로 바빠다’, ‘지금 막 퇴원했다’ 등) - 이런 반응자들에게 <변경

편지>를 보내고, Centre에서 가능하면 전화를 건다. 전화번호를 모를 때는 지속적인 협조를 바라는 글을 보낸다.

- 거부하기로 결정한 경우(‘다시는 재방문하지마시오’ 등) - 이번년도 조사에서는 어떠한 조치도 취하지 않는다. 그리고 다음년도 조사에서 재시도 해볼지 아닌지에 대한 판단은 Centre에서 한다.

차기연도 조사와 그 이후 조사에서 거부 전환을 위하여 ‘간단한 전화 면접’이 도입되었다. 전화 면접은 다음과 같은 집단에게 실시될 수 있다.

- 한 시간 면접동안 앉아 있을 수 없을 정도로 너무 아프거나, 나이가 많은 경우
- 최근 가족을 여의거나 이혼한 등의 개인적 상황이 있는 경우
- 누구를 돌봐야 해서 정말 시간이 없는 바쁜 경우

(4) 응답자와의 접촉 유지

가구패널조사에서 응답자와의 접촉을 유지하는 절차는 매년에 걸쳐 계속되는 중요하고 복잡한 일이다. Centre는 매년 조사를 위해 응답자의 주소 정보에 대한 광범위한 데이터베이스를 유지해야 하며, 최신 정보로 지속적으로 업데이트 해야 한다. 응답자와 적어도 2가지 유형의 의사소통이 조사연도 사이에 이루어지는데, 이는 9월 현장조사 전에 정확한 주거지 정보를 확인하기 위한 것과 응답자들 사이에 조사에 대한 인식을 심어주기 위해서 이다.

면접 이후에 Centre는 「감사 편지 thank-you letter」를 모든 피면접자에게 보내는데 이때 ‘선물 바우처’와 ‘주소변경 카드’(이사할 의도가 있을 때 공지)를 포함한다. 그리고 다음년차 조사 전에, 「주소확인 카드」와 함께 결과물에 대한 ‘요약 보고서’를 모든 성인(거부자 제외)에게 보낸다. 면접자는 모든 주소를 지속적으로 확인하여 현장조사 기간에 앞서 점검이 필요한 주소를 확인한다. 반송용 봉투를 동봉한 주소확인 카드를 통해 응답자가 자신의 주소를 확인하고 이사 나간 가구 구성원의 주소를 카드에 기록하여 통지하도록 요청한다. 이는 현장조사 동안 면접자가 해야 하는 추적 업무의 양을 최소화할 뿐 아니라, 각각의 주소에

대한 정확한 기록을 할 수 있도록 도와준다.

그리고 현장조사 전에 센터에서는 cover sheet에 위에 언급된 패널 유지 데이터베이스로부터 정보를 사용하여 표본 구성원의 최신 알려진 주소를 기입한다. 현장 조사 기간동안 응답자가 이사한 사실을 발견하게 되면 면접자는 다른 응답자, 새로운 세입자, 이웃으로부터 주소와 전화번호를 요구하고, 이 방식으로 주소를 알아 내지 못하면 지방 전화국, 상점, 우체국 등에 자문 요청을 하여 알아내도록 한다.

3) 독일 : The German Socio-Economic Panel Study(GSOEP)

가) GSOEP 개요

1984년 처음 실시된 GSOEP(Sozio-oekonomische Panel)은 독일의 대표적인 가구패널조사이다. 옛 서독지역에서 먼저 시작된 GSOEP은 같은 개별 가구, 개인, 가족들을 매년 조사하였고, 1990년부터는 옛 동독지역의 가구들을 포함하였고, 1994/95년에는 이민가구 역시 조사대상으로 포함되었다. 2003년에는 12,000가구의 약 24,000명 이상의 개인이 포함되었다.

GSOEP에는 개인과 가구에 대한 풍부한 자료가 포함되어 있는데, 예를 들면 객관적인 생활조건 및 주관적인 삶의 질, 다양한 생활영역의 변화 등, 이를 기초로 정치적·사회적 변화를 살펴보고 분석할 수 있다. 특히 매년 개인 및 가구의 소득통계를 산출할 수 있는 GSOEP 자료는 이를 토대로 개인뿐만 아니라 가구의 생활조건상의 안정성과 변화역시 분석할 수 있다.

GSOEP은 국립 경제연구기관인 DIW(Deutsch Institut für Wirtschaftsforschung)에서 전체를 관할하고, 설문지 작성 및 기타 실무는 민간학술조사 전문기관인 「Infratest」에서 수행하고 있다.²⁶⁾ 전체 연구 및 조사비용은 독일

26) 독일의 경우 조사 주관기관과 실사 수행기관의 신뢰성 문제를 극복하고 조사의 정도(精度)를 높이기 위하여 조사기록파일을 사용한다. 이 파일은 조사 기간 중에 나타난 여러 가지 특수한 상황이나 문제점을 기록한 파일로서 데이터의 신뢰성을 검토할 때 유용하게 사용할 수 있는 파일이며 부수적으로 조사방법을 연구하는 데에 귀중한 참고자료가 될 수도 있다.

국립과학재단인 DFG(Deutsche Forschungsgemeinschaft)에서 주로 지원을 하고 있으며 DIW로부터는 연구에 필요한 부대시설 및 필요연구원 인건비를 지원받고 있다.

GSOEP의 강점은 데이터 자체가 가지는 패널디자인, 모든 성인가구를 조사하는 가구의 맥락, 독일 내부 지역간의 비교, 현 독일 내에서 외국인에 대한 가장 규모가 큰 외국인 패널데이터²⁷⁾, 이민자에 대한 정보 등 이를 통한 연구의 가능성에 있다. 이를 기초로 GSOEP은 가구 구성, 직업 및 가족 일대기와 같은 광범위한 주제를 포괄하며, 조사 주제별 양식에서는 사회 보장, 교육 및 훈련, 시간 할당, 가족 및 사회 서비스 등의 주제를 포괄한다. 그리고 2000년부터 특히 아동 및 10대와 관련된 다양한 주제가 17세 가구 구성원을 위한 특별 설문지에 포함되었고, 2003년에는 아동 발달에 관한 질문들이 새로 태어난 아기를 가진 어머니들에 의해 수집되었다.

다음은 GSOEP 조사에서 매년 질문되는 본질적인 영역으로 비교적 안정적인 일련의 핵심 질문이다.

- 인구 특성(population and demography)
- 교육, 훈련, 자격증(education, training, and qualification)
- 노동 시장, 직업 변화(labour market and occupational dynamics)
- 수입, 소득, 사회보장(earnings, income and social security)
- 주거(housing)
- 건강(health)
- 가구 생산(household production)
- 기호 및 가치 등의 기본 지향(basic orientation)과 삶의 만족도 (satisfaction with life)

이러한 핵심 질문에 덧붙여 GSOEP은 조사 연도별로 주제를 선정하여 해당분야에 대한 정보를 수집하고 있는데, 그 주제들은 다음과 같다.

27) 현재는 가구주가 터키, 스페인, 이탈리아, 그리스 혹은 구 유고슬라비아출신일 경우에 한해서 조사되고 있음

〈표 2-19〉 조사 연도별 특별 주제

년 도	Wave	표본의 종류	주 제
1984	A / 1	A B	15세부터 고용 경력
1985	B / 2	A B	혼인 및 가족사
1986	C / 3	A B	사회적 배경, 첫 직업, 이웃 관계
1987	D / 4	A B	사회 보장, 조기 은퇴, 요보호 개인 및 자녀양육
1988	E / 5	A B	자산
1989	F / 6	A B	상급 교육이나 직업 훈련, 자격증
1990	G / 7	A B C	시간 활용과 기호 기초 질문(노동 시장 + 주관적 척도)
1991	H / 8	A B C	가족 및 사회적 서비스 가족 및 사회적 서비스(1984년 질문의 주관적 노동시장지수에 대한 반복질문 첨가)
1992	I / 9	A B C	사회보장과 빈곤(W4(1987)의 부분적 반복) 사회보장과 빈곤 (W4)의 노동시장 척도 및 일대기적 정보의 부분적 반복)
1993	J / 10	A B C	심화 교육이나 훈련(W6의 짧게 반복) 심화 교육이나 훈련, 노동 시장
1994	K / 11	A B C D1	이웃관계, 가치 및 기대 W11과 같고 더불어 이민 경력과 일대기
1995	L / 12	A B C D1 D2	W1의 부분적 반복 · 시간 활용과 기호, 추가된 임금 질문 W12와 같고 더불어 이민 경력과 일대기
1996	M / 13	A B C D	사회적 연결망 질문의 반복(W8)
1997	N / 14	A B C D	사회보장과 빈곤(W9의 반복)
1998	O / 15	A B D C E	생태학과 환경 행동(간접조세)
1999	P / 16	A B D C E	이웃 관계, 가치, 기대감
2000	Q / 17	A B D C E F	심화 교육, 훈련, 노동 시장
2001	R / 18	A B D C E F	사회적 연결망, 작업 조건
2002	S / 19	A B D C E F G	자산(W5 참조), 반복 사회보장(W14)
2003	T / 20	A B D C E F G	생태학과 환경 행동(W15)

주 : 표본의 종류에 대한 설명은 뒤에서 설명

자료 : Haisken-DeNew and J. R. Frick(Eds.) 2003.

이러한 GSOEP은 세계 최초의 가구패널조사인 미국의 PSID(Panel Study of Income Dynamics)를 모체로 하여 출발하였지만, 설문지 형식이나 조사 기법상 PSID의 문제점을 보완한 데이터라는 평가를 받고 있으며, 이후 유럽 각국에서 실시한 가구패널조사의 모태가 되었다.²⁸⁾

나) 표 본

(1) 표본 및 표집

1984년 초 조사를 시작하기 위하여, GSOEP 원 표본(original sample)은 1983년에 구 서독지역에서 선정되었고, 이후 1990년 6월부터 구 동독지역으로 확장되었다. 따라서 GSOEP에 의해 대표되는 모집단은 처음에는 서베를린을 포함한 서독지역의 거주 인구로 한정되었다가, 1990년 6월 이후 동베를린을 포함한 동독지역으로 확대됨으로써 독일 전체인구를 모집단으로 하는 독일의 대표적인 패널데이터로 자리잡게 되었다. 서독에서 선택된 외국인 집단들은 연구에서 과잉표집 되었고, 동독 지역 표본 비율이 서독의 표본비율보다 더 높다. 그리고 병원, 요양소, 군대와 같은 시설에 있는 사람들은 1차년도 조사에서 대표적으로 포함되지 못하였으나, 이후에 상기 종류의 시설에 일시적 혹은 영구적으로 거주하는 초기 가구의 사람들이 추적조사 되어졌다.

조사가 진행됨에 따라 GSOEP는 조사 지역과 모집단을 확장해 나갔는데, 다음과 같은 7개의 하위 표본을 가지며 각각은 다단계 무작위로 표집 되었다.

28) 강석훈(1997)

〈표 2-20〉 GSOEP 하위표본

하위	특징	시작연도	N (가구)	내용
A	West-German	1984	4,528	가구주 국적이 서독 혹은 이하 5개국 제외한 외국
B	Foreigners	1984	1,393 (과잉표본)	가구주 국적이 터키, 이탈리아, 스페인, 그리스, 유고슬라비아
C	East-Germans	1990	2,179	가구주가 국적이 동독(조사 영역 확장)
D	Immigrants	1994/ 1995	522	가구원 중 최소 한 사람이 1984년 이후 독일로 전입(조사 모집단의 확장)
E	Refreshment sample	1998	1,067	모든 현존 하위표본을 커버하는 무작위 표본(전체 모집단)
F	Innovation sample	2000	6,052	모든 현존 하위표본을 커버하는 무작위 표본(전체 모집단)
G	High Income Sample	2002	1,224	월 순가구수입 7,500 DM 이상

① Sample A “West German Residents”(“German Sample”)

표본 A는 가구주가 독일 국적을 가졌거나 혹은 터키, 그리스, 유고슬라비아, 스페인 및 이탈리아 국적이 아닌 다른 나라 국적을 가진 개별 가구에 사는 사람들을 포함한다. 소수 외국인만이 표본 A에 포함되기 때문에 이를 GSOEP의 ‘서독표본’이라고도 부른다. 표본 A의 표집은 1982년부터 「ADM(Arbeitsgemeinschaft Deutscher Marktforschungsinstitute)」원판에 기반을 두었고, 표집 설계는 ‘2단계와 1단계 내 2국면(two stages and two phases within the 1st stage)’을 갖는다. 1984년에 약 0.0002 표집확률로 4,528가구가 포함되어졌다.

② Sample B “Foreigners in West Germany”

표본 B는 1984년 옛 서독지역에 사는 터키, 그리스, 유고슬라비아, 스페인 및 이탈리아 국적의 가구주를 가진 개별 가구에 있는 사람들을 포함한다. 표본 B 모집단은 주(countries) 및 대도시(metropolitan) 지역의 PSUs(primary sampling units, 일차표집단위)로부터 선정되었고, PSUs의 무작위 선정은 각 국적별 독립적으로 추출되었다. 표본 A와 비교하여,

표본 B의 모집단은 과잉표집되었는데, 약 0.0002 표집 확률로 4,528 가구가 포함되어졌다.

③ Sample C "German Residents in the GDR(German Democratic Republic)"

1990년 봄 독일 통일이 기대되어졌고 GSOEP는 그해 6월에 GDR 영역으로 확장하였다. 표본 C는 이전 동독에 있는 개별 가구를 포함한다. 가구주가 GDR 시민인 개별 가구에 사는 사람들을 포함하는데 1990년 6월 GDR에 거주하는 모집단의 약 1.7%를 차지했다(외국인 제외). 「GDR중앙거주자파일(central resident' file of the GDR)」로부터 주소지를 접근하기 때문에 표본 A 및 B와는 다를 뿐 아니라 더 나은 표본 방법이 가능했다.

④ Sample D "Immigrants"

표본 D의 모집단은 표본 A, B, C를 선정한 모집단의 요소가 아니면서 적어도 한 명의 가구원이 1984년 이후 해외에서 독일로 이주한 가구로 정의된다. 표본 D는 1992/1994년에 선정된 D1과 1994/1995에 선정된 D2의 2개의 표본으로 구성되며 이들 각각은 다른 선정틀을 사용하였다. 첫 번째 표본인 D1은 236가구를, 두 번째 표본인 D2는 295 가구를 포함하였다.

⑤ Sample E "Refreshment"

1998년에 새로운 표본이 앞서 언급된 하위 모집단의 통합에 의해 주어진 독일에 사는 개별 가구의 모집단으로부터 선정되었다. ‘하위 표본 E라고도 표시되는 새로운 표본은 진행 중인 패널(A~D)로부터 독립적으로 선택되어졌다. 표본 E에 사용된 선택틀은 하위 표본 A를 선택하는데 사용된 틀과 본질적으로 유사하다. 2단계 표본의 선정상 약간의 차이가 있다 할지라도²⁹⁾, 하위 표본 A의 2단계 표본 선정할 때 사용한 것과 거의 유사하다.

29) 이는 주로 새로운 조사 도구, 예를 들면 CAPI(computer assisted personal interview) 대 PAPI(Paper-and-pencil personal interviews)에 의한 것이다.

⑥ Sample F "Innovation"

하위 표본 F는 2000년에 개별 가구 모집단으로부터 모든 다른 하위 표본으로부터 독립적으로 선택되어 졌다. 선정들은 응답 가구의 매달 월급이 최소 7,500 DM(EURO 3835) 이상인 것을 요구하였다.

(2) 표본 크기 확대(development)

GSOEP 조사는 응답률을 유지하기 위하여 공을 들였기 때문에 비교적 높은 안정성을 갖는다. 1984년 「GSOEP West」에 12,290명, 5,921가구가 참여하였고, 1990년 GDR에서 4,453명 2,179가구가 조사되었다. 이 표본은 「GSOEP East」 표본을 구성하였다. 최근 2002년 조사에서 「GSOEP West」 표본에는 7,175명 3,889 가구가, 「GSOEP East」 표본에는 3,466명, 1,818가구가 속하였다. 또한 1994년과 1995년의 522가구 1,078명의 이민자 표본에서도 유지율은 좋은 편이었다. 1998년 GSOEP은 보충적 표본 E의 1,067가구, 1933명 표본이 확대되었으며, 2000년에 GSOEP의 대규모 확장이 이루어 졌는데 표본 F는 6,052가구의 10,890명을 포함하였다(2002년; 4,586가구의 8,427명). 그리고 GSOEP은 2000년에 새롭게 고소득 가구(표본 G)의 무작위 표본 1,224가구, 2,671명을 보강하였다. 이러한 방법으로 표본이 확장되는 것을 통해 소규모 사회집단 분석의 가능성이 의미 있게 증가하였다.

이러한 GSOEP 표본 확장의 결정 요인은 (a) 인구학적 요인과 (b) 현장조사 관련 요인이 있다. 첫째, 인구학적 요인에는 개인의 탈퇴(사망, 해외로의 이사 등)와 진입(출생, 독일의 일부 지역 혹은 해외로부터 GSOEP 지역으로 전입, 최소 응답자 연령인 16세로 성장, 분가한 가구에 새로 들어온 가구 및 개인 등)과 관련된 요인이 있고, 둘째, 현장 조사와 관련된 요인에는 주어진 가구의 성공적인 접촉, 성공적인 면접 수행, 독신 가구나 부모 집을 떠난 젊은이 등 접촉하기 매우 어려운 사회적 집단 등과 관련된 요인이 있다.

다) 조 사

(1) 조사 원칙

GSOEP 조사의 원칙은 표준화된 조사 도구를 활용하여 직접 면접타 계식(face-to-face) 개인 면접을 가구 내 모든 16세 이상 개인에게 실시하는 것이다. 그리고 원칙적으로 대리 면접과 전화 면접은 하지 않으며 조사원에 의한 설문지 작성(paper-and-pencil)을 원칙으로 하며 가구 면접은 가구주를 대상으로 한다. 종단적 불일치를 줄이기 위해서 매년 가구설문지는 가구주가 가구를 떠나거나 사망하지 않는 한 그 사람에게 수행한다.

GSOEP 현장 조사가 가능한 한 구두 면접으로 수행되어야 함에도 불구하고, 부분적으로 응답자 스스로 수기하는 것이 허용되기도 한다. 그리고 최근에는 조사 방법의 혼합으로 면접자가 있는 상황에서 응답자가 스스로 작성하는 면접(self-administered interview)을 하거나, 면접자 없이 응답자가 완전히 작성한 다음 면접자를 찾으러 가거나 실조사기관(Infratest)에 우편으로 보내는 방법도 활용하고 있다.

그리고 응답자의 거부나 부재로 인해 면접이 성사되지 못할 경우엔 새로운 면접 날짜가 허용되며, 만약 응답자가 원할 경우 면접은 취소가 가능하며 예외적으로는 수기, 편지, 전화로 면접이 이루어 질 수도 있다. 미국 등지에서도 사용하고 있는 대리 면접은 GSOEP에서도 예외적인 사례에서는 가능하다. 그리고 1998년의 신 표본 E의 50%를 시작으로 CAPI(컴퓨터 지원 개인 면접)가 현장조사에 처음으로 도입되었다. 면접 방식의 변경 효과를 분석하기 위하여 표본 E 면접자의 나머지 50%는 기존 방식인 수기 면접(PAPI)으로 조사되었다. 추후 인터넷을 통한 자기 완성 면접(self-administerde interviewing via Internet) 방법을 도입하고자 2004년 현재 검토 중이다.

이러한 GSOEP 조사 규칙들은 융통성을 가질 수 있지만 응답자에 대한 정보는 반드시 응답자 본인을 통해 수집되어야 한다는 것만은 엄격하게 지켜져야 하는 규칙이다.

(2) 연대기적 질문

GSOEP은 1984년부터 개인 및 직업 일대기, 혼인과 청년 일대기, 노동시작 진입, 개인의 사회적 배경, 이민 정보 등 대표적인 일대기 정보가

수집되었다. 이 자료의 목적은 여성 노동 공급 측정 시의 출산 정보 등과 같은 많은 다양한 분석을 위해 중요한 기초 정보를 제공한다는 것과 직업력 또는 세대간 교육정도 등과 같은 독립 분석을 위해 정보를 제공한다는 것이다.

GSOEP은 다양한 도구와 설문지를 통해 조사에 참여하는 16세 이상 모든 개인의 일련의 회고적 일대기 자료를 수집하였다. 1984년에는 하위표본 A와 B를 대상으로 달력 양식의 직업력 조사에 초점을 두었고, 1985년에도 하위표본 A와 B를 대상으로 회고적 혼인 및 가족 일대기의 회고적 정보에 초점을 두었다. 1986년 설문지에서 사회적 배경과 직업 시장 진입을 다루는 일대기에 초점을 두었고, 이러한 질문 양식이 유지되었다. 그 후 1990년 처음으로 하위 표본 C(동독인)를 대상으로 정보를 수집하였고 1992년 동독인에게 직업, 교육 정도, 일대기 등과 관련된 다른 용어로 부가적 연대기 질문이 주어졌다. 서독인 설문지 버전과 비교하여 약간의 영역 확장이나 변경을 제외하고는 구조와 내용이 거의 동일하였다. 1994년에 별개의 일대기 설문지인 'Curriculum Vitae'가 4개의 하위표본(A, B, C, D1/D2)을 위해 만들어 졌고, 1996년부터 이것이 완전히 모든 하위 표본(A-D)을 위해 통합되어, 모든 질문이 모든 하위 집단에게 질문되었다. 그리고, 1999년 표본 E를 위한 일대기 설문지가 처음으로 도입되었다. 2000년에는 16세 혹은 17세 모든 응답자에게 해당하는 일대기 정보, 즉, 청소년(youth)에게 특별히 초점 맞추어진 설문지가 도입되었다. 2001년에는 일대기 설문지인 '생활사'(Lebenslauf)부분에서는 보다 심화되고 학교, 아동기 활동과 같은 보다 많은 질문이 포함되었다. 2002년에는 새로운 표본 G가 도입되었는데 이는 순가구 임금이 7,500 DM(EUR 3,850) 이상인 '부유한' 가구만을 표적으로 했다.

GSOEP의 질문은 시간에 따른 변화와 안정을 측정하기 위하여, 시간의 다른 측정(measurements-주어진 시기에서의 정보, 주기별 정보, 월별 정보, 삶의 일대기적 정보)을 사용하면서 시간의 다른 차원들(dimensions-과거, 현재, 미래)에 초점을 맞춘다. 조사 내용을 보면,

- 현재 시점에 대한 질문들(예: 최근 고용 상태, 최근 만족 수준)
- 수입에 대한 월별 달력과 노동 시장 이슈(예: 작년 1월부터 12월까지 고용상태)

- 일정 기간에 관련된 질문들(예: 작년 면접 이후 혼인, 배우자 사임과 같은 인구학적 변화)
- 미래 측면에 대한 질문들(예: 지금으로부터 향후 5년간 삶의 만족도 혹은 직업 기대도)

(3) 조사 도구

GSOEP의 조사도구는 가구에 대한 제반 정보가 포함되어 있는 Address Log와 설문지로 구성된다.

면접자가 작성하는 Address Log는 표지(cover sheet)라고도 불리며, 전반적인 가구정보를 포함하고 있다. 즉, 가구크기, 주거지역, 지역 정보와 가구원 개개인에 대한 성, 나이, 가구주와의 관계 등 가구 및 개인에 대한 전반적 정보, 현장조사 과정(연락 횟수, 탈락 사유, 면접 방법 등)에 대한 정보, 조사 상태에 따른 다른 버전의 설문지 표시(이전 주소 그대로인 원래 가구는 'green' 대 전출·분가 등 새로운 가구는 'blue'버전) 등의 정보를 포함하고 있어서 표본 탈락 분석, 표본 가중치, 종단면적 분석에 필수적이다.

GSOEP의 설문지는 경제학적 접근방법과 회학적 접근방법의 연계성을 강화하고, 가구라는 맥락 하에 개인의 경제활동을 파악하며 또한 횡단면적 분석과 종단면적 분석(longitudinal analysis)을 결합할 수 있도록 설계되었다. 매년 반복되는 주요 질문 주제로는 인구학적 변수, 노동 시장과 실업, 소득, 조세, 사회보장, 주거, 건강, 교육 및 직업훈련 등이 있다. GSOEP의 설문지는 가구용 설문지와 개인용 설문지, 비독일인 대상으로 하는 설문지 등으로 구분되어 있다. 각 설문지의 구체적인 내용은 다음과 같다.³⁰⁾

30) 면접자는 각 설문지를 완성한 뒤에는 면접 방법(구두, 직접 기록 등)과 면접 시간, 기타 코멘트, 일련 번호, 면접 수행 일자, 면접자 번호를 기록하고 사인을 하는 등 '면접자 평가'를 작성한다.

〈표 2-21〉 설문지의 내용

구 분	세 분	응 답 자
표준 도구	Address log	면접자 작성
	Household questionnaire	가구주
	Individual questionnaire	16세 이상
특정 집단 대상 부가 도구	life History	18세 이상 첫 응답자
	youth	16~17세 이상 첫 응답자
	mother & child	신생아 어머니
	gap	일시적 타락자

① 가구 설문지(Household Questionnaire)

가구 설문지는 가구주(Head of household)를 대상으로 실시하는데, 여기서 가구주라 함은 가구의 조건에 대해 가장 잘 아는 사람으로 정의되며 매년 조사대상이 된다. 내용은 다음과 같다.

- 주거 비용 질문: (공통) 이사 여부 및 이유, 주거지 특성, 소유 여부, 주거 크기, 만족도 등, (임차인만) 정부 주거 급여 수혜 여부, 주거 비용 등, (주택 소유자만) 대출 및 모기지 여부와 금액, 난방 비용, 주택 유지 비용 등, (공통) 저축이나 투자로 인한 수입 여부 및 액수, 유산·선물·복권 등으로 인해 2,500유로 이상의 큰 돈 획득 여부, 아동수당·가구 보조·환자보호 지원·사회부조 등의 급여 수혜 여부 및 금액, 가구원 전체 소득액, 자동차·PC·DVD기·인터넷 접속·핸드폰·팩스기 등의 소유 여부 등
- 기타 질문: 17세 이하 자녀의 성별 및 연령, 현재 자녀가 다니는 학교(기관), 자녀양육 지원자 여부 및 종류(친척, 공공 간호사, 사회복지사, 친구, 이웃 등), 주거지와 주요 시설들(쇼핑센터, 은행, 초등학교 등)과의 거리, 주거 지역의 환경 상태, 주거 지역의 범죄 수준이나 안전도, 외국인 거주 여부, 이웃과의 친밀도 등

② 개인 설문지(Individual Questionnaire)

개인 설문지는 조사대상 가구 내에 있는 16세 이상 개인에게 행해지는 설문지로 내용은 다음과 같다.

- 최근 생활 상황: 건강·직업·가정 임금·개인 임금·주거·여가 시간 등에 대한 만족도, 직업·쇼핑 등 가사 노동·자녀 양육·교육 및 훈련, 취미나 여가 활동 등에 사용하는 일주일간의 시간, 최근 교육 및 훈련 경험, 현재 고용 상태, 급여, 구직 활동 경험 및 이용 기관 등
- 현재 고용: 전직 시기, 이전 직업과 현재 직업 비교, 구직 방법, 직장 규모, 고용 기간, 직위, 주거지와 거리, 근로 시간, 연장 근무 수당 여부, 직종 종류, 급여, 연금·실업 급여·조기은퇴 급여 등 각종 급여 수혜 여부 및 금액 등
- 작년 활동: 지난 해 졸업 여부, 지난 해 취득 자격증 여부, 실업 이유 등
- 작년 수입: 수입원 및 금액, 피고용인 경우 급여 및 보너스 종류 및 금액, 은퇴자인 경우 연금 종류 및 금액 등
- 건강 및 질병: 현재 건강 상태, 스포츠, 체육관 경험, 흡연 여부, 신장 및 체중, 질병 유무, 의료적 재활 및 직업적 재활 여부, 건강보험 제공주체, 공적 및 사적 건강 보험, 사적 건강보험료 및 보장 범위 등
- 태도 및 의견: 정치 관심도, 지지 정당과 지지도, 사회문제 관심도 등
- 시민권과 국적: 독일 국적 여부, 독일 출생 여부, 시민권 종류, 비독일인 차별 경험 여부, 독일 체류 희망 기간 등
- 부가: 성별, 연령, 혼인 상태, 전년도 조사 이후 가족 상태 변화(혼인, 출산, 이혼) 등

③ 청소년 설문지(Youth Questionnaire)

2000년부터 아동 및 십대와 특별히 관련된 주제가 별도의 도구인 청소년용 설문지를 통해 GSOEP 데이터에 통합되었다. 보충적 청소년 설문지는 가구 내 개인 가운데 17~19세에 해당하는 청소년에게 실시하는 것으로 개인용 설문지를 보완하는 성격을 띤다. 이러한 십대에 대한 자료들은 아동기에 대한 회고적 정보가 부모의 인생 역사와 관련될 수 있기 때문에 이를 통한 다수의 분석을 가능하게 해 준다. 그리고 이 자료를 통해 세대간 분석 또한 가능해 진다. 구체적인 내용은 다음과 같다.

- 주거: 부모 집 거주 여부, 독립 주거나 개인 방 소유 여부 등

- 직업과 재정: 직업 유무, 취업 시작 연령 등
- 관계: 아버지, 어머니, 형제자매, 선생님, 친구, 또 다른 사람 등이 인생에서의 중요도, 그들과의 논쟁 여부 관계 등
- 스포츠와 여가: 여가 활동 종류 및 빈도, 스포츠 활동 장소 등
- 학교생활: 재학 여부, 최종 학력, 외국 재학 경험, 성적 만족도, 숙제 수행시 부모 도움 여부, 부모와 교사 연락 여부 등
- 직업 계획: 현재 직업 여부, 대학 진학 희망 여부, 경제적 자립 희망 연령, 희망 직업, 직업에서 중요 요소 등
- 장래: 대학 졸업 가능성, 구직 가능성, 실업 가능성, 자영업 가능성, 혼인가능성, 동거 가능성, 자녀 가능성 등 미래의 직업 및 개인 생활에 대한 가능성에 대한 질문 등
- 국적: 독일 출생 여부, 이민 연령, 이주전 거주 지역, 이주전 독일 내 친척 및 친구 존재 여부 등
- 아동기 및 부모 : 성장기를 보낸 지역 유형, 15세 이전 동거 가족원, 부모 출생연도, 부모의 독일 시민권 여부, 부모 직업상 지위 등
- 의견 및 신념 : 배우자 필요 여부, 자녀 필요 여부, 삶과자신에 대한 가치관 등

④ 보충적 일대기 설문지 (Supplementary Biography Questionnaire)

보충적 일대기 설문지(Lebenslauf)는 18세 이상인 가구 구성원에게 청년기와 초기 성인기에 대해 몇 가지 질문을 하는 것으로 개인 설문지 내용을 보완하는 것이다. 표본 F에 속하는 모든 남성과 모든 다른 하위 집단에 있는 새로운 응답자를 포함한다. 이것은 2001년부터 실시되었고 구체적인 내용은 다음과 같다.

- 시민권(Citizenship)과 국적(origin): 출생 연도, 독일 내 출생지, 국적, 독일 연방공화국으로의 전입 시기, 독일 시민권 소유 여부, 독일에 친지 및 친구 존재 여부 등
- 아동기와 부모 가정: 15세 이전 아동기 성장지역(대도시, 중소도시, 농촌 등), 15세까지 함께 산 가족구성원, 부모님의 독일시민권 소유 여부, 15세때 부모 직업 등

- 15세 이후 학교, 직업 훈련, 입대, 취업, 실업 등을 한 연령 등
- 교육: 최종 학력, 출신 학교 유형 재학 기간, 독일에서 직업 훈련/고등교육 수료 여부, 교육 종료연도 등
- 직업과 경력: 노동 시장 진입 연령, 직위, 고용 형태, 이직 경험, 자녀수, 강제 징병 혹은 방위 경력 등

⑤ 모자 설문지 Mother and Child Questionnaire

2003년에는 지난해에 새로 태어난 아기를 위한 설문지를 그들 엄마 대상으로 실시하였는데 이는 GSOEP 새로운 도구로 인생의 시작 시기에 있는 아동의 발달을 조사하고자 한 것이다. 조사 전년도 혹은 해당 연도에 출생한 아이에 대한 간단한 설문지로 개인적인 경험과 자녀 발달을 다룬다. 그렇기에 임신 상태도 포함하여 어머니만을 대상으로 하여 실시된다. 중심 주제는 아동 건강, 치료, 아동 발달, 부모 및 아동의 활동, 아버지 및 제3자의 아동 양육 및 지지 여부 등이다.

라) 관리체계

GSOEP의 표본관리체계에서 사망이나 해외 이주로 인해 조사에서 탈락되는 것은 자연스러운 이유이고 분석에서 큰 문제가 되지 않는다. 그러나 응답자의 거부나 가구 탈락(추후조사가 실패한 경우)으로 인한 패널 소실은 문제를 유발하는데 그것이 무작위가 아니기 때문이다.

GSOEP에서 유의미하게 중요한 패널 소실 원인은 첫째, '추후 조사 실패'로 가구가 이사한 경우, 분가(split-off)한 경우, 독신 가구 등이고 둘째, '응답 거부'는 동 베를린 거주, 가구주의 연령, 여성 가구주, 가구 이사, 분가, 별거 및 이혼, 면접자의 변화, 같은 면접자와의 면접 횟수, 낮은 가구 수입, 수입에 대한 항목 무반응, 동독에서 서독으로의 이주한 경우 등에서 빈번하게 일어난다.

상기 이유로 표본이 소실되는데 안정적인 패널 조사를 위해서 표본 구성원을 추적하고 지속적인 참여를 보장하는 것은 연구에서 매우 중요한 부분이라 할 수 있다.

(1) 추적(tracking)의 개념

전체 가구 구성원을 포함하는 패널 조사가 매년 대상자를 추적하는 기준은 선정된 목표 모집단의 대표성을 유지하는 것과 관련이 된다. 패널 조사는 이민을 제외하고 표적 모집단의 자연적인 인구 변화를 반영할 수 있다. 먼저 청소년 가구 구성원이 응답자 연령(16세)으로 성장하게 되면 조사의 대상이 된다. 둘째 ‘초기 가구’를 떠나는 개인은 특별한 패턴에 따라 추적되어진다. 이와 같이 종단적 분석 특히 인구학적 사건의 분석을 위해서는 초기 개인 뿐 아니라 GSOEP과 접촉했던 모든 개인들은 보다 잘 추적되어야 한다.

응답자의 이동성과 긴 패널 기간은 전체 표적 모집단에 영향을 주는데, 이보다 응답자의 연속적 조사에의 참여 거부가 표본 감소에 보다 큰 영향을 끼친다. 특히 이미 형성된 패널 가구로 전입하는 사람들이 조사를 거부하는 경우가 많다. 그리고 처음부터 참여한 개인은 표본 규모(size)에 중대한 문제를 일으키기 보다는 대표성(representativeness)에 영향을 끼친다. 그래서 이러한 요소들이 가중치 절차에서 고려되어야 하는 것이다.

GSOEP에서는 가능한 한 오래, 많은 응답자를 표집하기 위하여 표본 탈락에서 ‘최종 탈락’(final drop-outs)과 ‘일시적 탈락’(temporary drop-outs)을 구별한다. 집중적인 노력에도 불구하고 더 이상 개인이나 가구가 조사에 응하지 않는 경우 또는 추후 조사 참여에의 명백한 거부 의사를 밝힌 경우는 ‘최종’ 탈락이 된다. 그렇지 않은 경우는 처음에 ‘일시적’ 탈락으로 평가하고 추후 조사에서 가구에게 연락하고 참여를 설득하기 위한 새로운 시도를 한다. 이것이 실패하면 일시적 탈락은 최종 탈락이 된다.

(2) 패널 응답자의 동기 유지

응답자가 면접에 대한 동기를 갖는 것은 조사 참여에 중요한 영향을 미치는 요소이다. 패널 응답자가 조사에 참여하도록 동기를 유지하기 위한 방법은 다음과 같다.

첫째, 조사에 알기 쉬운 이름을 붙인다. 모든 표본 응답자들은 GSOEP의 주 내용을 ‘독일에서의 삶’(Life in Germany)으로 알고 있다.

둘째, 응답자는 GSOEP 연구의 목적에 대한 삽화가 그려진 「정보 안내서」를 제공 받는다(표본 B 경우엔 안내서가 응답자 각자의 모국어로 번역되어 있다).

셋째, 응답자는 데이터 비밀 보장에 대한 정보 용지(information sheet on data privacy)를 제공받는다.

넷째, 각 조사 마다 현장 조사 완수 시에는 「감사 편지(a letter of thanks)」를 보낸다.

다섯째, 각 응답자에게 잘 알려진 TV 「복권(lottery)」 티켓을 제공한다.

여섯째, 1987년부터(4th West-Wave) 모든 패널 가구에게 5~10 DM 상당의 「작은 선물(loyalty bonus)」을 제공한다(이는 서독에서 시간당 임금비율보다 적어야 한다).

덧붙여, 패널 조사에서는 응답자뿐 아니라 면접자들도 패널을 대표한다. 일부 면접자는 다른 면접자 보다 GSOEP이 요구하는 높은 응답률을 실현하기도 하고, 응답자 거부 가능성이 면접자에 따라 변화되기도 하였다. GSOEP은 면접자 배치 전략을 위해 가능한 한 많은 면접자를 조사 작업에 배치하고 면접자 수와 표본 지점(points) 수를 일치하도록 한다. 즉, 한 명의 면접자에 의해 집락(cluster)이 관리되도록 한다. 그리고 최소한 사례로 면접자가 조사 작업에 집중할 수 있도록 하는데 이러한 전략은 1984년 처음부터 사용하였다.

그리고 면접자의 동기 역시 응답자 조사 참여에 중요한 부분이다. 좋은 훈련, 프로젝트에 대해 충분한 정보, 조사 도구의 확실한 구조화, 조사 결과에 대한 정보 등은 성공적인 면접의 기초를 제공한다. 면접에 참여한 모든 면접자들은 매년 말에 조사 의뢰자인 DIW로부터 「감사 편지」를 받고, 면접자들은 요청에 의해서 GSOEP 결과를 기술한 책을 받을 수도 있다.

(3) 주소 업데이트

조사 기관은 Infratest에서 만드는 ‘패널 기록철’(master file)은 현장 조사 성공이 매우 중요한 요소이다. 이 파일에는 주소, 전화 번호, 면접 방법, 각 가구에 대한 여러 가지 정보들이 들어 있다. 조사 기관은 한 해 동안 면접 가구 및 개인의 주소가 변경되는지를 점검해야 한다. 조사 기

관의 주소 확인 작업은 조사 성공률에 주요한 영향을 끼치며 이때 주소는 우체국이나 지방거주자등록사무실 등을 통해 얻는다.

(4) 기 타

선물 등으로 응답자의 동기를 유지시키고, 주소를 업데이트 하는 방법 이외에 패널을 보호하기 위해서 GSOEP 데이터에 근거한 분석 결과가 담긴 소책자를 가구에 보내는 것도 중요하다. 그리고 GSOEP은 인간관계를 바탕으로 하는 「직접 대면 면접」 이고 좋은 관계가 유지되면 조사를 거절하기가 더 어려워지므로 시간 경과에 따른 면접자의 안정성이 매우 중요하다.

4) 국제비교 프로젝트

가) PACO 프로젝트

PACO 프로젝트는 Panel과 Comparability의 머리글자를 딴 이름으로 각국에서 실시되고 있는 패널 데이터를 국제비교학적인 관점에서 분석하기 위하여 데이터의 정의와 포맷 등을 일관되게 재정립하는 프로젝트이다. 이렇게 만들어진 PACO 데이터베이스를 이용하면 노동참가, 소득분배, 빈곤, 노인문제 등의 주제를 국제비교학적인 횡단면분석이나 패널 분석을 용이하게 수행할 수 있을 것이다.

PACO는 1992년 미국의 PSID, 독일의 GSOEP 그리고 룩셈부르크의 PSELL을 통합하여 데이터 포맷이나 데이터 정의 등에서 일관된 데이터베이스를 구축하는 프로젝트를 시작하였다. 1997년 자료에서 PACO Data Archive에는 유럽과 미국의 PSID 등 총 10개국의 데이터가 포함되어 있었다.

PACO 데이터베이스는 원패널 자료를 Compatibility와 Comparability의 관점에서 일관적으로 재정리한 자료이다. 각국에서 사용되는 여러 가지 개념이 유사성을 지니고 있으나, 국제비교학적인 관점에서 보면 국가간 상이한 요소가 많다. PACO 데이터베이스는 이러한 차이점을 일관성을

갖는 체계적인 분류 과정을 통해 국가간 상호조화되고 일관성을 갖는 변수와 파일들로 정리하여 얻은 결과이다. 현재 7개국의 데이터가 PACO 데이터베이스 안에 포함되어 있는데, 모두 관계형 데이터베이스 구조를 갖도록 설계되었으며, SPSS 파일로 동일한 변수이름, 값, 데이터 구조 등으로 이루어져 있다. 현재 사용가능한 PACO 데이터베이스 내용은 아래와 같다.

〈표 2-22〉 PACO 데이터베이스 구성내역

국 가	년 도	표본가구수/표본개인수
프랑스	1985~1990	2,100/7,500
독일	1984~1992	5,900/12,200
헝가리	1992~1994	2,100/5,800
룩셈부르크	1985~1992	2,000/6,000
폴란드	1987~1990	3,700/12,600
미국	1983~1987	6,800/19,400
영국	1991~1993	5,500/13,800

자료 : 강석훈, 「유럽의 패널조사 현황과 시사점」 재인용, p.23.

또한 PACO 데이터베이스의 내역으로는 소득변수 66개, 인구학적 변수 16개, 노동력 변수 29개, 인종학적 변수 14개, 주택변수 1개, 시간사용내역 4개, 가중치 4개, 교육 3개 등이다.

일관적이고 세밀한 방법으로 표준화된 PACO 데이터를 사용하면 다음과 같이 횡국가간(Cross Country) 패널데이터 분석을 용이하게 할 수 있다.

- 매크로 기능을 사용하면 PACO 데이터를 보다 쉽게 인출하거나 결합시킬 수 있다.
- PACO 데이터 구조는 전세계를 분석하는 프로그램을 이용할 수 있게 한다.
- 기본적인 SPSS(매크로) 프로그램을 수정하지 않고도 상이한 국가에서의 상이한 기간 간에 데이터 분석을 가능하게 한다.
- PACO 데이터를 분석하는 것이 원자료를 분석하는 것보다 용이하다.

- 연구자가 각국의 개별 패널데이터에 익숙하지 않아도 된다. 이와 함께 사용자는 PACO 데이터베이스에 포함되어 있지 않은 원자료를 PACO Data Archive에서 이용할 수도 있다.

나) LES(Luxembourg Employment Study) 프로젝트

LES(Luxembourg Employment Study)란 유럽연합의 Human Capital and Mobility Programme과 Nordic Research Council의 후원으로 이루어지는 국제노동연계 프로젝트이다. 알려진 바와 같이 1990년대에 접어들면서 많은 선진공업국가에서 특정 분야에서는 노동인력이 부족하면서, 다른 분야에서는 실업이 지속되는 노동시장의 불균형이 나타나게 되었다. LES는 이러한 현상을 국제비교학적인 관점에서 분석하는데 용이하도록 각국의 노동관련 데이터를 일관성 있는 체계로 재정리하는 프로그램이다. 뿐만 아니라 LES는 다양한 노동관련 국제비교연구를 수행하기 위하여 디자인되었는데 구체적인 연구목표는 다음을 포괄하고 있다.

- 취업과 실업의 국지적 패턴
- 취업과 실업에서의 변환성: 미국시장과 유럽시장의 비교
- 동부유럽 체제변환국의 노동시장 변화
- 인적자본과 경제적인 성과(economic performance)와의 관계분석
- 사회복지제도의 영향분석: 고용보험제도의 영향분석
- 직업탐색: 비경제활동인구(inactivity)와 실망근로자(discouragement)
- 이민/이주와 관련된 이슈
- 노인들의 노동시장행위: 은퇴결정
- 지역간, 직업간, 교육수준간, 성별, 연령별 실업의 본질 탐구

LES에는 1997년 기준 12개국의 고용관련데이터가 포함되어 있으며 향후 6~7개 국가가 더 참여할 것으로 예정되어 있다. 한편 이러한 각국의 데이터들은 1992년 EUROSTAT의 노동력과 관련된 집계변수에 사용된 분류체계와 ILO의 권고안, 그리고 EU의 국가들의 경험을 참조하여 아래의 12개의 주요그룹과 이에 관련된 90개의 변수로 구성되었다.

- 인구학적 변수들: 연령, 성별, 결혼상태, 응답자와의 관계, 국적, 가구원수와 가구내의 아이들의 수 및 취업자 수
- 노동현황: 설문주간에 근로여부 / 이유
- 주된 직장에서의 특성: 전문적인 지위, 산업, 직업, 근속기간 또는 직장의안정성, 근로시간
- 부업에 관한 정보: 산업, 취업, 근속기간 또는 직장의 안정성, 근로시간, 규칙성
- 과거의 취업경험: 전문적인 지위, 산업, 직업, 지나온 시간, 최종직장을 그만둔 이유
- 직장탐색: 기간, 주된 방법, 가용성, 등록여부
- 비경제활동인구의 상태
- 교육과 훈련: 최종 학력, 최종학력을 마칠 때의 연령, 교육훈련의 목적
- 조사 1년 전의 상태: 활동 종류, 전문적인 지위, 산업, 직업
- 노동시장에서의 상태: 취업/실업/비경제활동인구
- 수입과 소득
- 기술적인 항목들: 가구원수, 가중치 등

제 3 장 패널자료 분석방법론

패널데이터는 기본적인 N 명의 개인과 각각의 개인(i)에 대응되는 T_i 기간 동안의 관측치로 구성된 데이터이다. 이 때 모든 개인에 대하여 T_i 가 T 로 동일한 경우를 균형패널(balanced panel) 이라고 하고, T_i 각 개인별로 상이한 경우를 불균형패널(unbalanced panel)이라고 한다.

1. 분산분석(Analysis of Covariance)

일반적으로 패널데이터를 이용한 기본적인 선형 다변량분석모형을 다음과 같이 쓸 수 있다³¹⁾.

$$y_{it} = \alpha_i + x_{it}'\beta + \varepsilon_{it}$$

$$\text{단 } \beta' = (\beta_{1i}, \beta_{2i}, \dots, \beta_{Ki})', \quad x_{it}' = (x_{1it}, x_{2it}, \dots, x_{Kit}), \quad i=1, 2, \dots, N \\ t=1, 2, \dots, T$$

이 식의 경우에는 NT 개의 관측치로 $NT(K+1)$ 개의 파라미터를 추정하여야 한다는 점에서 추정가능한 형태는 아니다. 이에 따라 추정가능한 형태로서 다음과 같은 세가지 경우로 구분해 본다.

$$\langle 1-1 \rangle \quad y_{it} = \alpha_i + x_{it}'\beta_i + \varepsilon_{it}$$

$$\langle 1-2 \rangle \quad y_{it} = \alpha_i + x_{it}'\beta + \varepsilon_{it}$$

$$\langle 1-3 \rangle \quad y_{it} = \alpha + x_{it}'\beta + \varepsilon_{it}$$

<식 1-1>은 상수항과 기울기계수가 모든 개인에 따라 다르다는 형태이며, <식 1-2>는 상수항은 모든 개인에 따라 다르지만, 기울기계수는 모든 개인에 따라 동일하다는 형태이며, <식 1-3>은 상수항과 기울기가 모든 개인에 관계없이 동일하다는 형태이다. 이러한 세 가지 형태 중에서

31) 편의상 균형패널을 분석대상으로 한다. 불균형패널데이터를 이용하는 경우 선형분석 방법에서는 약간의 수정을 가하여 사용가능하다. 비선형분석의 경우에는 새로운 방법이 요구되는 경우가 많다.

선택하는 방법으로 다음과 같은 검정방법을 고려해 볼 수 있다.

가설 H_1 은 <식 1-1>에서 상수항은 상이하고, 기울기계수는 동일하다는 가설을 의미하고 H_3 은 <식 1-1>에서 상수항과 기울기계수가 모든 개인에 따라 동일하다는 점을 의미한다³²⁾.

가설 H_1 은 <식 1-1>에서 다음과 같은 $(N-1)K$ 개의 제약을 부과한 경우라고 볼 수 있다.

$$H_1 : \beta_1 = \beta_2 = \dots = \beta_N$$

또한 가설 H_3 은 <식 1-1>에서 다음과 같은 $(N-1)(K+1)$ 개의 제약을 부과한 경우라고 볼 수 있다.

$$H_3 : \alpha_1 = \alpha_2 = \dots = \alpha_N, \beta_1 = \beta_2 = \dots = \beta_N$$

가설 H_3 을 검정하기 위해서는 정규분포의 가정 하에서의 다음의 F-통계량을 사용할 수 있다. 이 때 F-분포의 자유도는 각각 $(N-1)(K+1), NT - N(K+1)$ 이다.

$$F_3 = \frac{(S_3 - S_1) / [(N-1)(K+1)]}{S_1 / [NT - N(K+1)]}$$

여기에서 $S_1 = \sum_{i=1}^N RSS_i$ (<식 1-1>을 모든 개인에 대하여 회귀분석할 때의 잔차제곱합의 총합)이다. 즉,

$$RSS_i = W_{yy,i} - W_{xy,i} W_{xx,i}^{-1} W_{xy,i}$$

$$W_{yy,i} = \sum_{t=1}^T (y_{it} - \bar{y}_i)^2, \quad W_{xx,i} = \sum_{t=1}^T (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)',$$

$$W_{xy,i} = \sum_{t=1}^T (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i) \quad \bar{y}_i = (1/T) \sum_{t=1}^T y_{it} \quad \bar{x}_i = (1/T) \sum_{t=1}^T x_{it}$$

$$S_3 = T_{yy} - T_{xy} T_{xx}^{-1} T_{xy}' \quad (\text{<식 1-3>의 잔차제곱합})$$

32) 이외에도 상수항은 동일하지만 기울기가 개인에 따라 다르다는 가설도 고려해 볼 수 있지만 이는 현실적으로는 의미가 없는 가설이다.

$$\begin{aligned}
 T_{yy} &= (1/NT) \sum_{i=1}^N \sum_{t=1}^T (y_{it} - \bar{y})^2, \\
 T_{xx} &= (1/NT) \sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x})(x_{it} - \bar{x})', \\
 T_{xy} &= (1/NT) \sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x})(y_{it} - \bar{y})', \\
 \bar{y} &= (1/NT) \sum_{i=1}^N \sum_{t=1}^T y_{it} \quad \bar{x} = (1/NT) \sum_{i=1}^N \sum_{t=1}^T x_{it}
 \end{aligned}$$

만약 가설 H_3 가 채택되면 자료를 통합(pool)하여 <식 1-3>의 단일 방정식을 추정한다. 만약 H_3 가 기각되면 H_1 가설을 검정한다.

가설 H_1 을 검정하기 위해서는 다음과 같은 F-통계량을 사용한다. 이때 F-분포의 자유도는 각각 $(N-1)K, NT-N(K+1)$ 가 된다.

$$F_2 = \frac{(S_2 - S_1)/(N-1)K}{S_1/[NT-N(K+1)]}$$

여기에서 $S_2 = W_{yy} - W_{xy}W_{xx}^{-1}W_{xy}$ (<식 1-2>의 잔차제곱합).

$$W_{xx} = \sum_{i=1}^N W_{xx,i} \quad W_{yy} = \sum_{i=1}^N W_{yy,i} \quad W_{xy} = \sum_{i=1}^N W_{xy,i}$$

만약 H_1 이 기각되면 <식 1-1>을 이용한다. 만약 H_1 이 채택되면 <식 1-2>를 이용한다³³⁾.

2. 정태 선형분석(Analysis of Static Linear)

우리는 다음과 같은 회귀분석 모형을 고려한다.

$$\langle 2-1 \rangle \quad y_{it} = \alpha_i + x_{it}'\beta + \varepsilon_{it} \quad i=1,2,\dots,N \quad t=1,2,\dots,T$$

33) 이와 같은 방법은 추정계수가 시간에 따라 변화한다는 모형에서도 동일하게 사용될 수 있다. 즉, 다음과 같은 모형에도 적용된다.

$$\begin{aligned}
 y_{it} &= \alpha_t + x_{it}'\beta_t + \varepsilon_{it} \\
 y_{it} &= \alpha_t + x_{it}'\beta + \varepsilon_{it} \\
 y_{it} &= \alpha + x_{it}'\beta + \varepsilon_{it}
 \end{aligned}$$

가. 고정효과모형(Fixed Effects Model)

고정효과모형은 α_i 를 고정된 값으로 취급하는 모형이며, 최소자승더미변수모형(least squares dummy-variable model, LDV) 또는 공분산모형(analysis of covariance model)이라고 불린다.

<식2-1>는 다음과 같이 벡터형태로 쓸 수 있다.

$$\text{단, } Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} e \\ 0 \\ \vdots \\ 0 \end{bmatrix} \alpha_1 + \begin{bmatrix} 0 \\ e \\ \vdots \\ 0 \end{bmatrix} \alpha_2 + \dots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ e \end{bmatrix} \alpha_N + \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \beta + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{bmatrix}$$

$$y_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{bmatrix}, \quad X_i = \begin{bmatrix} x_{1i1} & x_{2i1} & \dots & x_{ki1} \\ x_{1i2} & x_{2i2} & \dots & x_{ki2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1iT} & x_{2iT} & \dots & x_{kiT} \end{bmatrix}, \quad e = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \quad \varepsilon_i = \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \vdots \\ \varepsilon_{iT} \end{bmatrix}$$

$$E(u_i) = 0, \quad E(\varepsilon_i \varepsilon_i') = \sigma^2 I_T, \quad E(\varepsilon_i \varepsilon_j') = 0 \quad \text{if } i \neq j$$

잔차제곱합을 최소화하는 방법으로 다음과 같은 추정량을 얻을 수 있다.

$$\hat{a} = \bar{y}_i - \bar{x}_i' \hat{\beta}_{CV}$$

$$\hat{\beta}_{CV} = \left[\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' \right]^{-1} \left[\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i) \right]$$

$$\text{단, } \bar{x}_i = \frac{1}{T} \sum_{t=1}^T x_{it}, \quad \bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it}$$

$\hat{\beta}_{CV}$ 은 집단내추정량(Within Estimator), 공분산추정량(Covariance Estimator) 그리고 고정효과추정량(Fixed Effect Estimator)을 불린다.

$\hat{\beta}_{CV}$ 은 다음과 같은 형태로 표기될 수 있다.

$$\hat{\beta}_{CV} = \left[\sum_{i=1}^N X_i' Q^{-1} X_i \right]^{-1} \left[\sum_{i=1}^N X_i' Q^{-1} y_i \right]$$

$$\text{단, } Q = I_T - (1/T)ee'$$

공분산추정량은 다음과 같은 성질을 가진다.

$$\begin{aligned} \text{Var}(\hat{\beta}_{CV}) &= \sigma_\varepsilon^2 \left[\sum_{i=1}^N X_i' Q X_i \right]^{-1} \\ \text{plim}_{N \rightarrow \infty} \hat{\beta}_{CV} &= \beta \\ \text{plim}_{T \rightarrow \infty} \hat{\beta}_{CV} &= \beta \\ \text{plim}_{T \rightarrow \infty} \hat{a} &= a \\ \text{plim}_{N \rightarrow \infty} \hat{a} &\neq a \end{aligned}$$

$\hat{\beta}_{CV}$ 는 최소분산불편선형추정량(BLUE)가 된다. 상수항의 경우에는 T 가 증가하면 모수에 대한 일치추정량이 되지만, N 이 증가하는 경우에는 일치추정량이 되지 않는다. 동 추정량은 기본적으로 개인효과를 차분을 통해 제거하는 방법을 사용하고 있기 때문에 추정모형에 시간불변 변수(time-invariant regressors)가 있는 경우에 동 변수의 계수값은 추정되지 않는다.

나. 임의효과 모형

<식 2-1>는 다음과 같이 쓸 수 있다.

$$\begin{aligned} y_{it} &= \mu + x_{it}'\beta + \alpha_i + \varepsilon_{it} = \mu + x_{it}'\beta + u_{it} \\ \text{단, } u_{it} &= \alpha_i + \varepsilon_{it} \end{aligned}$$

개별 효과를 나타내는 α_i 를 확률변수로 취급하는 동 모형은 임의효과모형(Random Effect Model) 또는 분산구성모형(Variance Component Model) 또는 일방향오차구성모형(One-way Error Component Model)으로 불린다.

이 모형에서는 이제 오차항 u_{it} 가 두개의 구성요소를 가지게 되는데, 각각에 대하여 다음을 가정하자.

$$E(\alpha_i) = 0$$

$$\begin{aligned}
 E(\varepsilon_{it}) &= 0 \\
 E(\alpha_{it}) &= \sigma_\alpha^2 \text{ if } i=j, \quad E(\alpha_{it}) = 0 \text{ if } i \neq j \\
 E(\varepsilon_{it} \varepsilon_{js}) &= \sigma_\varepsilon^2 \text{ if } i=j \text{ and } t=s, \quad E(\varepsilon_{it} \varepsilon_{js}) = 0 \text{ of} \\
 &\text{otherwise} \\
 E(x_{it} \alpha_{jt}) &= E(u_{it} x_{it}) = 0
 \end{aligned}$$

이 때 오차항의 분산은 다음과 같다.

$$E(u_i u_i') = \sigma_\varepsilon^2 I_T + \sigma_\alpha^2 e e' = V, \quad V^{-1} = \frac{1}{\sigma_\varepsilon^2} \left[I_T - \frac{\sigma_\alpha^2}{\sigma_\varepsilon^2 + T \sigma_\alpha^2} e e' \right]$$

예를 들어 $T=3$ 인 경우 이 분산행렬은 다음과 같다.

$$V = \begin{bmatrix} \sigma_\varepsilon^2 + \sigma_\alpha^2 & \sigma_\alpha^2 & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\varepsilon^2 + \sigma_\alpha^2 & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\alpha^2 & \sigma_\varepsilon^2 + \sigma_\alpha^2 \end{bmatrix}$$

임의효과모형을 공분산추정량으로 추정하는 경우에도 동 추정량은 기울기계수에 대하여 불편추정량이고 일치추정량이 되지만, 효율적인 추정량은 되지 못한다. 왜냐하면 공분산추정량은 그룹내 변동(within group variation)만을 반영할 뿐 그룹간 변동(between group variation)을 반영하지 못하기 때문이다.

이제 오차항의 이분산성을 감안한 일반최소자승추정량은 다음과 같다.

$$\langle 2.2 \rangle \quad \hat{\delta}_{GLS} = \left[\sum_{i=1}^N \tilde{X}_i' V^{-1} \tilde{X}_i \right]^{-1} \left[\sum_{i=1}^N \tilde{X}_i' V^{-1} y_i \right]$$

단, $\delta' = (\mu, \beta')$. 이 일반최소자승추정량은 추정하고자 하는 모수의 최소분산선형불편추정량이 된다.

한편, 동 일반자승최소추정량은 다음과 같이 도출될 수도 있다. 먼저

$$\bar{y}_i \text{ 와 } \bar{x}_i \text{ 의 } (1-\Psi^{1/2}) \text{ 부분만을 차감한다(단, } \Psi = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T\sigma_\alpha^2}).$$

그리고 차감된 즉, $[y_{it} - (1-\Psi^{1/2})\bar{y}_i]$ 를 상수항과 $[x_{it} - (1-\Psi^{1/2})\bar{x}_i]$ 에 대하여 회귀분석한다. 이 때의 추정량은 일반최소자승추정량과 같게 된다.

한편, σ_α^2 와 σ_ε^2 가 미지의 모수이기 때문에 실제 일반최소자승추정량을

사용하려면 동 파라미터를 추정해야 한다. 먼저, 1단계에서는 공분산추정량과 같은 일치추정량을 이용하여 동 파라미터들을 추정한 다음 2단계에서 동 추정치를 대입한 분산행렬을 이용하여 일반최소자승추정량을 구한다. 적절한 표본크기 하에서($T \geq 3$, $N - (K + 1) \geq 9$, for $T = 2$, $N - (K + 1) \geq 10$) 이러한 2단계 추정량은 공분산추정량보다 효율적이다.

다. 그룹내 그리고 그룹간 추정량(Within and Between Estimator)

우리는 다음과 같이 세가지 방법으로 통합회귀분석(pooled regression)을 고려할 수 있다.

$$\langle 2-3 \rangle \quad y_{it} = \alpha + x_{it}'\beta + \varepsilon_{it}$$

$$\langle 2-4 \rangle \quad \overline{y_{it} - y_i} = (\overline{x_{it} - x_i})'\beta + \varepsilon_{it} - \varepsilon_i$$

$$\langle 2-5 \rangle \quad \overline{y_i} = \alpha + \overline{x_i}'\beta + \varepsilon_i$$

OLS방법을 <식 2-3>, <식 2-1-7>과 <식 2-8>에 적용하여 통상적인 OLS추정량 $\hat{\beta}_{OLS}$ 그룹내 추정량 $\hat{\beta}_W$ 그리고 그룹간 추정량 $\hat{\beta}_B$ 을 얻을 수 있다. 약간의 변형을 통하여 우리는 다음과 같이 OLS추정량이 그룹내 추정량과 그룹간 추정량의 가중평균임을 보일 수 있다.

$$\hat{\beta}_{OLS} = F^W \hat{\beta}_W + (I - F^W) \hat{\beta}_B$$

$$\text{단,} \quad F^W = [S_{xx}^W + S_{xx}^B]^{-1} S_{xx}^W, \quad S_{xx}^W = \sum_{i=1}^N \sum_{t=1}^T (x_{it} - \overline{x_i})(x_{it} - \overline{x_i})'$$

$$S_{xx}^B = \sum_{i=1}^N T(\overline{x_i} - \overline{x})(\overline{x_i} - \overline{x})'$$

또한 GLS추정량도 다음과 같이 그룹내 추정량과 그룹간 추정량의 가중평균임을 보일 수 있다.

$$\hat{\beta}_{GLS} = H^W \hat{\beta}_W + (I - H^W) \hat{\beta}_B$$

단, $H^W = [S_{xx}^W + \Psi S_{xx}^B]^{-1} S_{xx}^W$.
 만약 $T \rightarrow \infty, \Psi \rightarrow 0$ 가 되고, 이 때 $\hat{\beta}_{\alpha S} = \hat{\beta}_W = \hat{\beta}_{\alpha V}$ 이 성립한다.

라. 고정효과 또는 임의효과 모형의 선택

고정효과모형은 개인간의 차이가 회귀분석의 대상이 되는 파라메터의 변화(parametric shifts)로 확신할 수 있을 때 합리적인 접근방법이다. 이 모형은 표본내의 개인에게 적용되는 것으로 이해할 수 있으며, 표본 이외에는 동 모형이 적용되지 않는다는 점을 가정하고 있다. 예를 들러 국가가간 비교에서 모든 국가가 포함되어 있는 경우라면 개별효과가 고정되어 있다고 보는 것이 합리적일 것이다.

이에 반해 개별표본들이 이들을 포함하는 더욱 큰 모집단에서 임의로 추출된 표본이라고 보는 경우에는 개별효과가 임의적으로 분포되어 있다고 판단하는 것이 합리적일 것이다. 이러한 경우에는 PSID 또는 KLIPS와 같은 경우에 해당된다.

실제 실용성의 측면에서만 판단한다면 더미변수접근 방법은 횡단면 표본수가 많은 중단면자료의 경우에 자유도의 측면에서 우위가 없어지며, 이 때에는 임의효과모형을 사용하는 것이 직관적으로 선호된다. 하지만 고정효과 모형도 큰 장점을 가지고 있다. 즉, 임의효과모형에서와 같이 개별효과가 다른 설명 변수들과는 모형에 사용하는 다른 설명변수와 상관관계가 없다는 가정이 합리성을 찾기가 쉽지 않다. 따라서 고정효과 모형은 생략변수에서 오는 추정량으로 불일치성에서 자유롭기 어렵다.

고정효과와 임의효과 모형의 구분은 T 가 많아지면 공분산추정량과 일반최소자승추정량이 같아지기 때문에 의미가 없어진다. T 가 고정되어 있고, N 이 큰 경우 고정이나 임의의 선택은 쉽지 않다. 실제로 두 모형의 추정치는 매우 큰 차이를 보인다. 많은 개인에 대해서 적은 수의 관측치만 있는 경우에는 적은 수의 관측치로부터의 정보를 효과적으로 사용하는 것이 효율적인 추정량을 얻는데 매우 중요해진다.

고정효과모형은 연구자가 관계성을 표본내로 한정하는 조건부 분석(conditional inference)을 실시하는 것으로 볼 수 있다. 이에 반해 임의효

과모형은 연구자가 모집단 전체에 대하여 비조건부 분석(unconditional inference) 또한 한계분석(marginal inference)을 실시하는 것을 볼 수 있다. 실제로 조건부 분석과 비조건부 분석의 선택은 연구자의 몫이라고 하겠다. 실제로 연구결과를 표본내로 한정한다면 고정효과 모형이 적절하며, 연구결과가 표본의 행태분석결과를 이용하여 모집단의 행태분석에 위한 것이라면 임의효과 모형이 적절하다고 하겠다.

마. Hausman 검정

임의효과모형에서 임의효과를 나타내는 항과 독립변수간의 부상관성을 검정하는 방법으로 Hausman검정을 실시할 수 있다. 동 검정의 기본 아이디어는 상관관계가 없다는 가설 하에서는 공분산추정량과 일반최소자승추정량이 모두 일치추정량이지만 공분산추정량이 비효율적이며, 대립가설에서는 공분산추정량은 일치성을 가지지만, 일반최소자승추정량은 일치성을 갖지 못한다는 점에 근거한다. 구체적인 검정 통계량은 다음과 같다.

$$[\hat{\beta}_{CV} - \hat{\beta}_{GLS}]' \Sigma^{-1} [\hat{\beta}_{CV} - \hat{\beta}_{GLS}]$$

단, $\Sigma = \text{Var}(\hat{\beta}_{CV}) - \text{Var}(\hat{\beta}_{GLS})$. 상관관계가 없다는 귀무가설 하에서 위의 검정통계량은 점근적으로 자유도가 K 인 카이제곱 분포를 따르게 된다.

바. 개인효과와 시간효과를 모두 포함하는 모형

고정효과모형은 다음과 같이 쓸 수 있다.

$$y_{it} = \alpha_i + \gamma_t + x_{it}'\beta + \varepsilon_{it} \quad \text{단, } i=1,2,\dots,N \quad t=1,2,\dots,T-1$$

$$y_{it} = \mu + \alpha_i + \gamma_t + x_{it}'\beta + \varepsilon_{it}$$

$$\text{단, } i=1,2,\dots,N \quad t=1,2,\dots,T-1 \quad \text{and} \quad \sum_{i=1}^N \alpha_i = 0, \quad \sum_{t=1}^T \gamma_t = 0$$

기울기계수의 최소자승추정량 $\hat{\beta}_{CV}$ 은 종속변수를 $(y_{it} - \bar{y}_i - \bar{y}_t + \bar{y})$, 독립변수를 $(x_{it} - \bar{x}_i - \bar{x}_t + \bar{x})$ 로 사용하는 회귀분석에서 얻을 수 있다.

단, $\bar{y}_t = (1/N) \sum_{i=1}^N y_{it}$.

임의효과모형은 α_i 와 γ_t 가 모두 확률변수이고, x_{it} 와 상관관계가 없다면 다음의 가중치행렬을 사용하는 일반최소자승추정량이 최소분산선형불편추정량이 된다.

$$V = \sigma_\varepsilon^2 I_{NT} + \sigma_\alpha^2 I_N \otimes e e' + \sigma_\gamma^2 e_N e_N' \otimes I_T$$

단, $\sigma_\gamma^2 = \text{Var}(\gamma_t)$.

사. 불균형 패널

고정효과모형인 경우에는 약간의 변형을 통해 추정가능하다. σ_ε^2 , $\text{Var}(\hat{\beta}_{CV})$ 와 F-통계량의 추정시에 자유도의 조정이 필요하며, 개별 하위의 평균을 구할 때 표본수를 조정하면 된다.

임의효과모형인 경우 분산행렬의 대각원소가 더 이상 동일한 크기가 아니기 때문에 더 이상 $I \otimes V$ 가 되지 않는다. 또한 \sqrt{t} 가 T 에 의존하기 때문에 이분산성의 문제가 발생된다. GLS를 일부분 보완하여 사용하면 추정이 가능한데, 균형패널에서 사용하였던 Ψ 대신에 다음과 같은 Ψ_i 를 사용하여 추정을 실시한다.

$$\Psi_i = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T_i \sigma_\alpha^2}$$

3. 동태 선형분석(Analysis of Dynamic Linear)

개별효과를 포함하는 동태패널분석에서는 고정효과모형이던 임의효과모형이던간에 공분산추정량은 불일치하게 된다. 다음의 모형을 고려해 보자.

<3-1> $y_{it} = \theta y_{i,t-1} + \alpha_i + \varepsilon_{it}$

단, $i=1,2,\dots,N$ $t=1,2,\dots,T$, $|\theta|<1$ 이고 y_{i0} 는 관측가능하다. 공분산추정량은 다음과 같이 쓸 수 있다.

$$\langle 3-2 \rangle \quad \hat{\theta}_{CV} = \frac{\sum_{i=1}^N \sum_{t=1}^T (y_{it} - \bar{y}_i)(y_{i,t-1} - \bar{y}_{i,-1})}{\sum_{i=1}^N \sum_{t=1}^T (y_{i,t-1} - \bar{y}_{i,-1})}$$

<식 3-2>의 공분산은 다음과 같은 편향(bias)을 가진다.

$$p \lim_{N \rightarrow \infty} (\hat{\theta}_{CV} - \theta) = -\frac{1+\theta}{T-1} \left(1 - \frac{1}{T} \frac{1-\theta^T}{1-\theta}\right) \left[1 - \frac{2\theta}{(1-\theta)(T-1)} \left(1 - \frac{1-\theta^T}{T(1-\theta)}\right)\right]^{-1}$$

이러한 편향이 발생하는 이유는 공분산추정량에서 개별 효과를 제거하는 과정에서 제거한 모형에서의 잔차와 설명변수가 상관관계를 가지기 때문이다. 또한 임의효과모형에서는 최소자승추정량은 모수를 과대추정한다(정확한 과대추정의 정도는 Hsiao, p.73를 참조하라).

동태모형의 추정에는 MLE(Maximum Likelihood Estimator), GLS(Generalized Least Squares Estimator), IV(Instrumental Variable Estimator), GMM(Generalized Method of Moments) 등이 사용된다.

가. 임의효과 모형의 추정

MLE, GLS, IV, GMM 등의 방법이 가능하다. MLE의 경우 최초 조건을 잘못 선택하면 일치성을 보장받지 못한다. IV형태의 추정량은 최초 조건으로부터 자유롭다. IV추정량의 기본 아이디어는 다음과 같다. <식 2-9>에서 다음을 얻을 수 있다.

$$\Delta y_{it} = y_{it} - y_{i,t-1} = \theta \Delta y_{i,t-1} + \Delta \varepsilon_{it} \quad \text{where } t=2,3,\dots,T$$

이 때 다음과 같은 적률조건(moment condition)을 얻을 수 있다.

$$\begin{aligned} E(y_{i,t-2} \Delta \varepsilon_{it}) &= 0 \\ E(y_{i,t-2} \Delta y_{i,t-1}) &= 0 \end{aligned}$$

따라서 우리는 $y_{i,t-2}$ 를 IV로 이용하여 Θ 를 추정할 수 있다. 또한 다음과 같은 적률조건도 구할 수 있다.

$$\begin{aligned} E(y_{i,t-2} \Delta \varepsilon_{it}) &= 0 \\ E(y_{i,t-3} \Delta \varepsilon_{it}) &= 0 \\ &\vdots \\ E(y_{i,0} \Delta \varepsilon_{it}) &= 0 \\ E(\Delta y_{i,t-2} \Delta \varepsilon_{it}) &= 0 \\ E(\Delta y_{i,t-3} \Delta \varepsilon_{it}) &= 0 \\ &\vdots \end{aligned}$$

이러한 적률조건들을 이용하여 GMM 추정을 할 수 있다. 독립적인 설명변수가 포함되어 있는 경우에는 다음과 같은 Arellano and Bond의 GMM 추정량을 사용할 수 있다. 단, $\delta = (\Theta, \beta')$.

$$\begin{aligned} y_{it} &= \Theta y_{i,t-1} + x_{it}' \beta + a_i + \varepsilon_{it} \\ \hat{\delta}_{GMM,AB} &= \left\{ \left[\sum_{i=1}^N \begin{pmatrix} \Delta y_{i,t-1} \\ \Delta x_{i,t-1}' \end{pmatrix} W_i' \right] \left[\sum_{i=1}^N W_i \bar{A} W_i' \right] \left[\sum_{i=1}^N W_i (\Delta y_{i,t-1}, \Delta x_{i,t-1}') \right] \right\}^{-1} \\ &\times \left\{ \left[\sum_{i=1}^N \begin{pmatrix} \Delta y_{i,t-1} \\ \Delta x_{i,t-1}' \end{pmatrix} W_i' \right] \left[\sum_{i=1}^N W_i \bar{A} W_i' \right]^{-1} \left[\sum_{i=1}^N W_i \Delta y_i \right] \right\} \end{aligned}$$

단, $q_{it} = (y_{i,0}, y_{i,1}, \dots, y_{i,t-2}, x_i')$, $x_i' = (x_{i1}', x_{i2}', \dots, x_{iT}')$

$$W_i = \begin{pmatrix} q_{i2} & 0 & \cdots & 0 \\ 0 & q_{i3} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & q_{iT} \end{pmatrix}, \quad \bar{A} = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 2 \end{pmatrix}$$

GMM 추정량은 (y_i, x_i') 가 i 에 대해 iid 인 경우에는 a_i 가 고정이던 확률변수이던 간에 관계없이 일치성을 가지며, 점근적 정규분포를 가진다.

나. 고정효과 모형의 추정

공분산추정량과 정규분포MLE는 일치성을 보장받지 못한다. GMM 추정량은 (y_i, x_i') 가 i 에 대해 iid 인 경우에는 a_i 가 고정이던 확률변수이

던 간에 관계없이 일치성을 가지며, 점근적 정규분포를 가진다. 이 경우에는 또한 변환우도비추정량(transformed likelihood estimator)을 사용할 수 있다. 변화의 모형과 측정 방법(2002 한국청소년패널조사 194~212)

4. 패널 분석의 예³⁴⁾

다음은 농가경제조사 패널화 자료를 이용하여 정태선형회귀분석을 실시한 예를 보여주고 있다. 패널분석을 하려면 먼저 기본적인 데이터 형태를 설정하여야 한다. 일반적으로 패널자료를 정렬하는 방법은 long form과 wide form으로 구분할 수 있다. 예를 들어 변수 var에 대하여 2003년부터 3년간 관측한 자료가 있는 경우를 상정해 보자. 이 때 데이터는 다음과 같이 정렬되어 있을 때 long form이라고 한다.

<표 3-1> long form 자료

i	j			
id	year	var	sex	...
1	2003	3,000	1	
1	2004	4,000	1	
1	2005	5,000	1	
2	2003	2,300	0	
2	2004	2,350	0	
2	2005	2,500	0	
...		

또한 다음과 같은 형태로 정렬되어 있을 때 wide form이라고 한다.

34) 본고에서는 STATA를 이용하여 패널분석을 실시한 예를 보여주고 있다. SAS, SPSS등의 통계패키지를 이용하여 패널분석을 실시할 수도 있다.

〈표 3-2〉 wide form 자료

i				
id	sex	var2003	var2004	var2005
1	1	3,000	4,000	5,000
2	0	2,300	2,350	2,500
...				

STATA에서는 reshape이라는 명령어를 통해 wide form과 long form의 데이터를 상호 변환할 수 있다. 다음의 이러한 변환의 예를 보여주고 있다.

```
reshape wide g08 g17 g18 g04 mage f07 f08 f09 f12 f13 f14 c16 c17 c19 c21,
i(rootkey) j(year) (note: j = 2003 2004 2005 2006 2007)
```

```
Data                                long  ->  wide
-----
Number of obs.                       20350 ->  4070
Number of variables                    18   ->   77
j variable (5 values)                 year  -> (dropped)
xij variables:
g08  ->  g082003 g082004 ... g082007
g17  ->  g172003 g172004 ... g172007
g18  ->  g182003 g182004 ... g182007
g04  ->  g042003 g042004 ... g042007
mage ->  mage2003 mage2004 ... mage2007
f07  ->  f072003 f072004 ... f072007
f08  ->  f082003 f082004 ... f082007
f09  ->  f092003 f092004 ... f092007
f12  ->  f122003 f122004 ... f122007
f13  ->  f132003 f132004 ... f132007
f14  ->  f142003 f142004 ... f142007
```

104 농가소득의 동태적 변화요인

```

c16  ->  c162003 c162004 ... c162007
c17  ->  c172003 c172004 ... c172007
c19  ->  c192003 c192004 ... c192007
c21  ->  c212003 c212004 ... c212007

```

```

. sum

```

Variable	Obs	Mean	Std. Dev.	Min	Max
rootkey	4070	3.46e+07	3312920	1.10e+07	3.92e+07
c162003	3197	2.305286	10.92193	-6	205
c172003	3197	.7719737	7.076968	0	137
c192003	3197	4.551454	23.76057	-77	716
c212003	3197	50.06788	1306.346	-7	57000
....					
c162004	3196	2.094493	10.28829	0	231
c172004	3196	.6864831	6.173289	0	98
c192004	3196	4.600751	24.95428	0	799
c212004	3196	75.39956	1949.78	-13	88300
....					
c162005	3198	2.198874	10.39543	0	199
c172005	3198	.69606	6.368679	0	105
c192005	3198	4.427142	22.37758	-3	617
c212005	3198	51.0738	1052.584	-2	43135
....					
c162007	3185	2.186185	10.90843	0	207
172007	3185	.7494505	6.655939	0	139
c192007	3185	4.519623	24.43549	0	576
c212007	3185	49.8697	1125.337	-9	40979
....					
sum_y	4070	3.769287	1.81474	0	5

이제 이 데이터를 long form으로 변환하여 보자.

```
. reshape long g08 g17 g18 g04 mage f07 f08 f09 f12 f13 f14 c16 c17 c19 c21,
    i(rootkey) j(year) (note: j = 2003 2004 2005 2006 2007)
```

Data	wide	->	long
Number of obs.	4070	->	20350
Number of variables	77	->	18
j variable (5 values)		->	year
xij variables:			
g082003 g082004	g082007	->	g08
g172003 g172004	g172007	->	g17
g182003 g182004	g182007	->	g18
g042003 g042004	g042007	->	g04
mage2003 mage2004 ...	mage2007	->	mage
f072003 f072004	f072007	->	f07
f082003 f082004	f082007	->	f08
f092003 f092004	f092007	->	f09
f122003 f122004	f122007	->	f12
f132003 f132004	f132007	->	f13
f142003 f142004	f142007	->	f14
c162003 c162004	c162007	->	c16
c172003 c172004	c172007	->	c17
c192003 c192004	c192007	->	c19
c212003 c212004	c212007	->	c21

```
. sum
```

Variable	Obs	Mean	Std. Dev.	Min	Max
rootkey	20350	3.46e+07	3312594	1.10e+07	3.92e+07
year	20350	2005	1.414248	2003	2007

c16		15974	2.199136	10.71923	-6	238
c17		15974	.7148491	6.540973	0	155
c19		15974	4.524352	24.05068	-77	799
-----+						
c21		15974	55.04188	1337.356	-13	88300
f07		15974	1498.806	2435.382	0	60287
f08		15974	788.5776	1504.351	0	35164
f09		15974	381.3199	1175.901	0	19663
f12		15974	1333.491	3364.598	0	91202
-----+						
f13		15974	978.9337	2983.821	0	81840
f14		15974	104.107	698.4801	0	22235
g04		15985	1522.746	1579.453	0	31436
g08		15985	1.21e+07	2.22e+07	-3.94e+08	7.79e+08
g17		15985	3.01e+08	3.57e+08	-1.17e+08	1.03e+10
-----+						
g18		15985	3.11e+07	6.37e+07	0	3.00e+09
mage		15985	60.74758	10.56106	28	89
sum_y		20350	3.769287	1.814562	0	5

이제 long form으로 정렬된 데이터를 이용하여 정태선형패널회귀분석을 실시한다. 다음은 임의효과(random effect, re)모형과 고정효과(fixed effect, fe) 모형의 예를 각각 보여주고 있다.

```
. xtreg g08 g17 g18 g04 mage f07 f08 f09 f12 f13 f14 c16 c17 c19 c21, re
Random-effects GLS regression           Number of obs   =   15963
Group variable (i): rootkey            Number of groups =   3740

R-sq:  within = 0.0193                  Obs per group:  min =    1
        between = 0.5019                  avg   =    4.3
        overall = 0.3110                  max   =    5

Random effects u_i ~ Gaussian           Wald chi2(14)    =  4289.01
```

corr(u_i, X) = 0 (assumed) Prob > chi2 = 0.0000

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
g08						
g17	.0059245	.0005221	11.35	0.000	.0049012	.0069479
g18	-.0146611	.0029911	-4.90	0.000	-.0205235	-.0087986
g04	4208.87	133.5796	31.51	0.000	3947.059	4470.681
mage	-17625.45	18906.44	-0.93	0.351	-54681.39	19430.49
f07	1420.007	80.91753	17.55	0.000	1261.411	1578.602
f08	1072.54	127.9128	8.38	0.000	821.8354	1323.244
f09	1519.206	167.3098	9.08	0.000	1191.285	1847.128
f12	702.4493	58.41032	12.03	0.000	587.9672	816.9314
f13	177.8677	66.67395	2.67	0.008	47.18912	308.5462
f14	1714.151	254.1063	6.75	0.000	1216.112	2212.19
c16	267516.2	17969.15	14.89	0.000	232297.4	302735.1
c17	608329.5	29625.49	20.53	0.000	550264.6	666394.4
c19	19744.19	7729.981	2.55	0.011	4593.709	34894.68
c21	67.33868	123.8797	0.54	0.587	-175.461	310.1384
_cons	-576427.4	1248433	-0.46	0.644	-302331	1870457
sigma_u	8226507.1					
sigma_e	16609086					
rho	.19699593 (fraction of variance due to u_i)					

. xtreg g08 g17 g18 g04 mage f07 f08 f09 f12 f13 f14 c16 c17 c19 c21, fe

Fixed-effects (within) regression Number of obs = 15963

Group variable (i): rootkey Number of groups = 3740

R-sq: within = 0.0283 Obs per group: min = 1

제 4 장 농가경제조사 자료의 사후 패널화

1. 농가경제조사의 개요³⁵⁾

가. 조사개요

1) 조사목적

경제사회 변천에 따른 농가경제의 동향과 농업경영실태를 파악하여 농업정책 수립과 농업경영 개선을 위한 기초 자료를 제공하기 위해 실시하는 조사이며, 농가소비지출, 노동투하량 등 각종 농가경제지표를 생산하여 농업문제 연구를 위한 실증적 자료를 제공한다.

2) 조사연혁

1953년 농림부와 한국은행이 합동으로 「농촌실태조사」를 처음 실시하였으며, 1954년 농림부가 「농촌실태조사」를 「농가경제조사」와 「농산물 생산비조사」로 분리하여 1962년 통계법에 의한 정부 공식통계로 지정하였다. 1962년부터 1998년까지 7차에 걸쳐 표본을 개편하였고, 1998년 7월 1일 정부조직법 개정에 따라 통계청으로 이관하여 2003년 표본개편(8차)과 함께 조사표 및 조사지침을 개편하였다.

3) 조사주기 및 기간

조사주기는 매월이며 조사대상기간은 매년 1월 1일 ~ 12월 31일이다.

4) 조사대상

전국 320개 표본 조사구내 3,200농가를 조사대상으로 한다.

35) 통계행정편람 2007, 통계청 참조

<표 4-1> 조사대상 비교표

조사대상 가구	제 외 가 구
<ul style="list-style-type: none"> · 10a(1,000m²)이상의 경지를 직접 경작하는 가구 · 연간 농축산물의 판매액이 50만원 이상으로 농업을 계속하는 가구 - 단, 판매금액이 50만원 미만이라도 조사 기준시점 현재 50만원 이상의 가축을 사육하는 가구는 포함 	<ul style="list-style-type: none"> · 단독가구(1인 가구) 및 외국인 가구 · 비혈연가구원으로 구성된 가구 · 상시고용인 5인이상의 대규모 기업농 · 농업시험장, 학교, 종교단체 및 법인 등의 준농가

주 : 「2005농림어업총조사」 농가구분 정의 적용

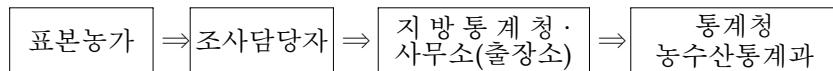
5) 조사방법

<표 4-2> 일계부와 원부 조사방법

조사표 종류	조 사 사 항	조사시기	조사방법
일 계 부	<ul style="list-style-type: none"> ① 작물재배현황 ② 도정량 ③ 수입·지출 ④ 농업노동 투입내역 ⑤ 농산물생산비 투입내역 ⑥ 자가농업생산물 중 자가소비량 	매 월	농가에서 직접 기입하거나, 조사담당자에 의한 면접조사
농 가 원부	<ul style="list-style-type: none"> ① 농가자산 ② 농가부채 	연초, 연말 및 수시	조사담당자에 의한 면접조사

- 일 계 부 : 표본농가에 일계부를 비치하여 매일 매일의 농업경영 및 가계운영에 따른 수입·지출내역을 기입하며, 농업노동 투입내역 및 농산물생산비 투입내역은 조사담당자가 직접 면접 조사한다.
- 농가원부 : 표본농가의 자산상태를 파악하기 위하여 연 2회(자산 변동시는 수시 조사) 농가의 재산 상태를 경영주와의 면접을 통하여 조사한다.

6) 조사체계



7) 조사 결과 공표

다음 해 3월경 조사 결과를 요약하여 보도 자료로 공표한다. 공표이후 4월경 『농가경제통계』 연보에 전국평균, 전·겸업별, 주·부업별, 영농형태별, 경지규모별, 경영주연령별, 가구원규모별, 지역별 등으로 자료를 구분하여 수록·발간한다.

보도 자료는 통계청 인터넷 홈페이지(www.nso.go.kr)에, 연보내용은 국가통계 포털(KOSIS)에 게재한다.

8) 법적근거

통계법 제 8조에 의한 지정통계 제 10142호이다

나. 표본설계

1) 표본설계 연혁

농가경제조사(표본조사)는 '62년부터 시작하여 45여년간 지속되어 오고 있으며, 농업여건의 변화를 반영하기 위해 5년마다 표본을 재설계하고 있다.

2002년 한국통계학회에 의뢰하여 2000 농업총조사 결과를 토대로 설계하였다.

2) 모 집 단

2000농업총조사 결과 조사된 1,384천 농가를 전체 모집단으로 하고 도별자료의 생산을 위하여 각 도의 농가를 부차 모집단으로 하였다.

〈표 4-3〉 모집단(농가수) 변화추이

(단위 : 천가구, %)

	1950	1960	1970	1980	1990	2000	2005
전국가구	3,791	4,371	5,576	7,969	11,355	14,312	15,887
농가	2,474	2,350	2,587	2,155	1,767	1,384	1,273
(비율)	65.3	53.8	46.4	27.0	15.5	9.7	8.0

3) 표본설계 내용

가) 층화

현행 표본설계의 층화방법을 따르되 논벼 층을 해당 조사구의 논벼 농가 비율에 따라서 70%이상 층과 50~70%층으로 구분하고 해당 영농 형태 층으로 분류하기 위한 최소기준(10%이상)을 주는 방법이다.

나) 층화방법

영농형태를 기준으로 층화함으로써 통계조사 결과의 연속성을 유지한다. 논벼 층을 논비율에 따라서 세분하여 층화함으로써 표본의 모집단에 대한 대표성을 높였고, 표본의 구성이 논벼 비율에 대해서 안정적이다. 최소한의 기준(해당 영농형태의 농가 비율이 10% 이상)을 설정함으로써 논벼 비율과 2종 겸업농가 비율이 각각 40% 이상인 경우에는 극히 소수 가구의 영농형태에 따라서 구분되게 되는 문제점을 개선한 것이다.

다) 표본 추출틀(조사모집단)

단독농가와 농가규모가 20호 미만인 조사구는 모집단에서 제외한다. 조사구 : 23,268개(농가 : 1,115,252호)이다.

층화기준은 다음과 같다.

〈표 4-4〉 표본추출틀의 층화기준

구 분	층 화 기 준
논벼지구1	전업 또는 1종 겸업농가이면서 논벼 농가의 비율이 70%이상인 지구
논벼지구2	전업 또는 1종 겸업농가이면서 논벼 농가의 비율이 50%-70%인 지구
과수지구	논벼 및 2종 겸업농가의 비율이 50%이하이고, 과수 농가가 최대인 지구(10% 이상)
채소지구	논벼 및 2종 겸업농가의 비율이 50%이하이고, 채소 농가가 최대인 지구(10% 이상)
특작지구	논벼 및 2종 겸업농가의 비율이 50%이하이고, 특작 농가가 최대인 지구(10% 이상)
화훼지구	논벼 및 2종 겸업농가의 비율이 50%이하이고, 화훼 농가가 최대인 지구(10% 이상)
전작지구	논벼 및 2종 겸업농가의 비율이 50%이하이고, 전작 농가가 최대인 지구(10% 이상)
축산지구	논벼 및 2종 겸업농가의 비율이 50%이하이고, 축산 농가가 최대인 지구(10% 이상)
기타지구	위의 영농특성 층이나 2종겸업 층에 속하지 않는 지구
2종겸업	2종 겸업농가의 비율이 50%이상인 지구

라) 표본배정법 결정

선택기준은 전국통계치에 대한 상대표준오차와 각 영농형태에 대한 기대농가수 및 영농형태별 통계의 상대표준오차 그리고 각 지역별 통계에 대한 상대표준오차 등을 기준으로 한다.

표본배정법은 절충배정법으로 농가소득, 농업소득, 농가부채의 세 변수에 대한 네이만배정에 따른 표본크기의 평균으로 결정한다.

마) 지역별 표본수 및 주산지 결정

각 도에 대해서 30개의 조사구를 우선 배정하고, 나머지 표본은 비례 배정법에 따라서 추가로 배정한다. 해당 농작물의 주산지 시·군에 대해서 조사구별 해당 농작물의 재배면적과 해당 농가비율을 계산하여 주산지에 대한 표본 조사구를 결정한다.

바) 각 지역의 층별 표본배분

농가소득, 농업소득, 농가부채 변수에 대해서 각각 네이만배분법에 따라 층별 표본수를 구한 후에 절충하는 절충배정법을 이용하고 있다.

사) 표본추출

층화와 표본배정을 마친 후 각 지역에서 층별로 배정된 표본 조사구 수만큼을 조사구 내의 농가수에 비례하는 확률비례계통추출법을 적용하여 추출한다.

아) 추출단위

1차(PSU)은 농업조사구 또는 부락한다. 시부지역은 농업조사구, 군부지역은 누락 또는 리 단위로 한다. 단 제주도는 군부와 시부지역 모두 농업조사구로하여 추출한다. 2차(SSU)은 농가단위로 추출한다.

자) 목표오차

농가경제에 관한 각종 지표 중 가장 중요하고 다른 변수들과 상관관계가 높은 농가소득을 통제변수로 삼아 농가소득의 전국평균에 대한 추정량의 표본오차(표본변이계수를 사용)가 2% 이내가 되도록 설계하였다.

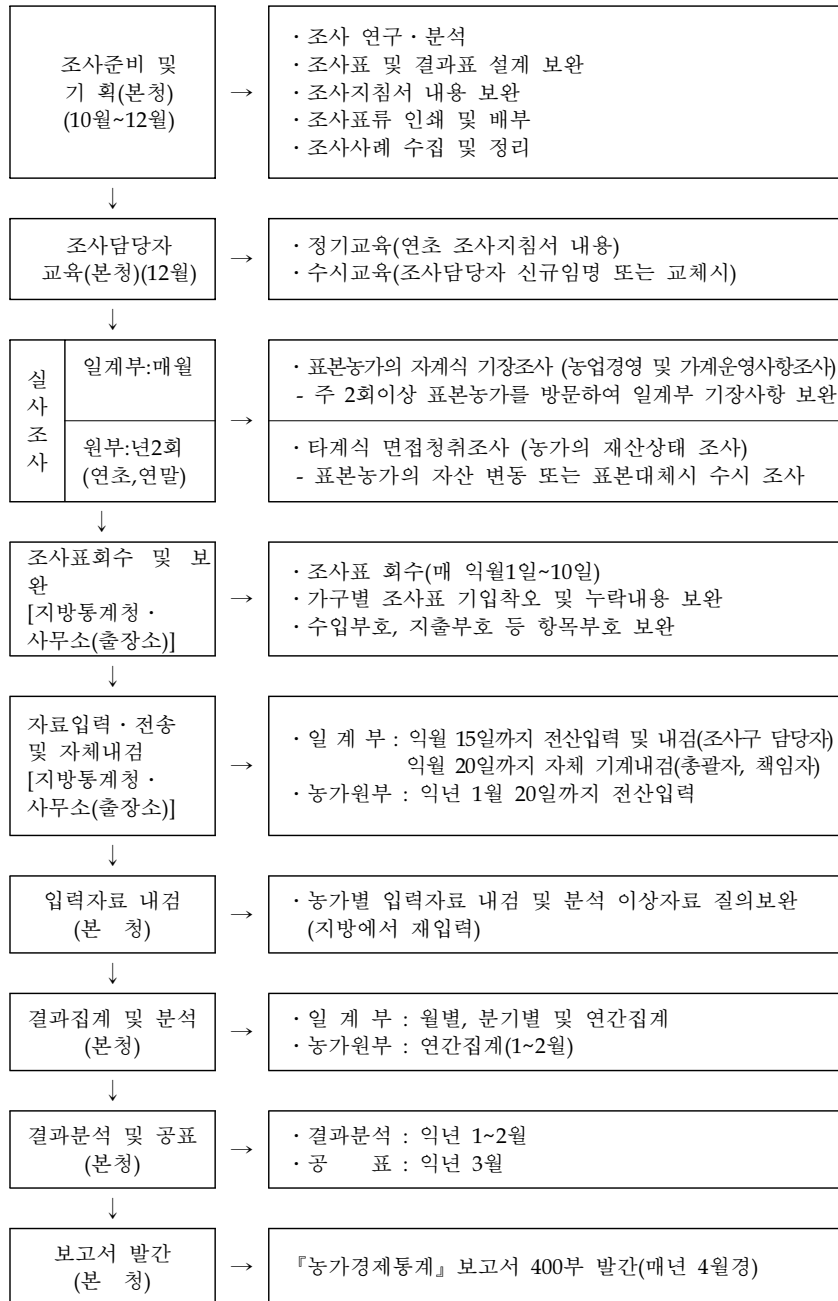
모집단의 농가수 감소로 인하여 표본오차(조사결과 측정된 변이계수)는 점차 줄어들게 된다.

4) 추 정

층화2단 추출법에 의해서 표본 농가를 선정하여 조사한다. 층화는 전국의 9개 각 지역에 대해서 영농형태에 따라서 10개 층으로 구분한다. 1차 추출단위는 조사구이고, 2차 추출단위는 각 농가이다.

추정내용은 전국 농가소득, 농업소득, 농가부채 등의 추정하고, 각

라. 농가경제조사 업무흐름도



다. 농가경제 수입 및 지출 항목분류

1) 수입 분류

사업 수입	농업수입	작물수입	미곡, 맥류, 잡곡, 두류, 서류, 채소, 특작, 과일, 화훼, 기타작물, 농작물 부산물
		축산수입	대동물, 소동물, 축산물, 축산부산물
		농업잡수입	농업소득피해보상금, 폐농자재 판매대금 등
	겸업수입	임업 및 어업수입, 제조업수입, 건설업수입, 기타겸업수입	
사업외 수입	근로수입	급여수입, 상여금수입, 농업임금수입	
	자본수입	이자 및 배당수입, 임대수입, 유가증권 매매차익	
이전 수입	공적보조	농업보조금, 기타공적보조금	
	사적보조	출타가족보조, 친인척등 보조	
재산적 수입	자산의 감소	고정자산의 매각, 유동자산의 감소	
	부채의 증가	부채의 차입, 미불금 및 선수금의 발생	
	비경상수입	경조수입, 퇴직일시금, 재산수증, 사고보상금, 가사잡수입	

2) 지출 분류

사업지출	농업지출	생산원가	① 재료비(종자 및 종묘, 비료, 농약, 소동물, 사료, 양잠 기타양축, 기타재료) ② 노무비(지불임금) ③ 경비(영농광열비, 수선비, 농구비, 이자비, 위탁수수료, 임차료, 水利費, 조세부담금, 이자비용, 영농잡지출)
		판매 및 관리비	도정료, 농업보험료, 판매수수료, 판매자재비, 생산관리비
	겸업지출	임업 및 어업 비용, 제조업 비용, 건설업 비용, 기타 겸업 비용 겸업용 부채의 이자비용	
사업외 지출	농외취업 비용, 자본수입을 위한비용(유가증권매매손실 포함) 농외 차입금에 대한 이자		
조세부담금	< 해당 부문별로 항목 재분류 >		
가계지출	소비지출	식료품비, 주거비, 광열수도비, 가구집기·가사용품, 피복신발비, 보건의료비, 교육비, 교양오락비, 교통통신비, 기타소비지출	
	비소비지출	가계용 조세·부담금, 가계부문 차입금 이자, 송금 및보조(출타자녀 학비)	
재산적 지출	자산의 증가	고정자산의 구입, 유동자산의 증가	
	부채의 감소	부채의 상환, 미불금 및 선수금의 정산	

바. 농가재산 분류 및 정의

재산의 분류			정 의	
농 가 재 산	고 정 자 산	토 지	○ 농업경영 및 기타용으로 이용되는 소유지로서 대부지를 포함	
		건 축 물	건축물	○ 농업 및 기타용으로 이용되는 주택, 창고, 헛간, 축사, 퇴비사, 온실 등의 지상건축물 중 취득가격이 10만원 이상의 것(취득가격 100만원 이상의 농업 사업전용건축물을 포함)
			구축물	○ 과수원, 액비통, 사이로, 우물, 수문, 용수로, 명거, 압거, 등의 토지개발시설과 다른 구축물 중 취득가격이 10만원 이상의 것
		유 형 자 산	대 농 구	○ 농업용으로 사용하는 기계기구 중 취득가격이 10만원 이상의 것(집합농구를 제외)
			기계· 기구 및 비품	집합농구 ○ 농가가 사용함에 있어 통상 수개 내지 수십개를 동시에 사용함으로써 그 목적을 달성할 수 있는 농구로서 취득가격이 10만원 이상인 것(양계용케이지, 육묘상자 및 농산물수확상자, 누에솥 등)
			자동차등	○ 농업 및 기타용으로 제공되는 오토바이, 스쿠터(배기량 50cc이하 포함), 삼륜자동차, 승용차, 트럭, 라이트밴 및 지프 중 취득가격이 10만원 이상인 것(취득가격 100만원 이상의 농업사업전용기계 포함)
		비 품	○ 농업용 및 기타용으로 이용되는 PC, FAX, 복사기 중 취득가격이 10만원 이상인 비품(취득가격 100만원 이상의 농업사업전용비품 포함)	
		대 식 물	○ 농업용으로 사용되며 감가상각 대상이 되는 영년성작물 ※ 정원 및 택지에서 분산 재배하고 있는 과수는 자산에 포함하지 않음.	
		대 동 물	○ 소, 말, 돼지, 염소, 꿀벌 및 마리당 가격이 10만원(어미기준)이상인 동물	
		무 형 자 산	○ 경작권, 어업권, 영업권, 전세권 등	
	유 동 자 산	재 고 자 산	소 동 물	○ 가금을 비롯하여 위 대동물에 해당되지 않는 모든 동물 ※ 단, 애완용동물은 제외
			미처분 농축산물	○ 처분하지 않고 재고로 가지고 있는 농업생산물(가정용 또는 농업, 기타용으로 사용예정인 것을 포함).
			미사용 구입자재	○ 농업용으로 구입한 원료 및 보조원료. 구체적으로는 종묘, 비료, 사료, 농약, 제제료 등
당 좌 자 산		현 금	○ 보유 현금	
		예 금 등	○ 예금, 계 및 저축성보험금, 대부금, 주식, 공·사채, 투자신탁, 기타 유가증권 등.	
미수금 및 선급금	○ 농산물의 외상매출금, 기타 미수입금, 구매계약 등과 관련하여 미리 지급한 돈			
부 채	차 입 금	○ 정부, 각종단체, 지방공공단체, 농협, 은행, 거래처, 개인 등으로 부터의 차입금		
	미불금 및 선수금	○ 농업생산자재나 기계용품 등의 외상매입 미불금과 계약선수금		

사. 농가의 농업노동 분류 및 정의

노동의 분류		정 의	
농 가 노 동	직 접 노 동	경 중 부 문	묘상작업 ○ 모판이나 온상에서 (묘)를 키워서 본답에 옮겨 심거나 판매하기 위해 묘를 키우는데 투입한 작업일체
		본답준비 ○ 작물을 파종하거나 옮겨심기 위하여 본답(논, 밭)을 갈거나 고르는 작업(로타리작업 포함) 일체	
		파종작업 ○ 본답에 직접 씨앗을 뿌리거나 종자를 심는 작업과 종자준비작업	
		이식작업 ○ 묘(묘)를 모판에서 본답으로 옮겨 심는 작업과 보식작업	
		본답관리 ○ 작물의 생육기간 중에 성장과 결실을 돕기 위해 본답에서 수행하는 작업중 시비·제초·병충해 방제 이외의 모든 작업	
		비료주기 ○ 본답에 무기질비료나 유기질비료를 운반하여 뿌리는 작업	
		제초작업 ○ 본답의 잡초제거를 위한 일체의 작업	
		병해충방제 ○ 병충해 예방 및 구제를 위해 본답에서 수행하는 일체의 작업	
		수확작업 ○ 최종 생산물을 거두어 들이기 위하여 작물을 베거나 뽑는 작업, 알뿌리를 캐는작업, 열매를 따는 작업, 탈곡 전의 건조작업 및 묶어 세우는 작업 포함(수확시 선별작업 포함)	
		수확물운반 ○ 수확한 농작물을 탈곡장, 건조장 또는 최종 보관장소까지 옮기는 작업	
		건조작업 ○ 수확직후에 생산물의 수분함량을 적정수준으로 낮추기 위하여 말리는 작업	
		축 산 부 문	먹이주기 ○ 사료급여·조제 및 급수작업
	방역및위생관리 ○ 가축 및 축사의 위생관리를 위한 방역, 목욕시키기 등		
	청소및퇴비처리 ○ 축사청소, 깔집반입, 퇴비처리 등		
	기타사육관리 ○ 착유, 교미, 운동시켜주기 등		
	간 접 노 동	출하준비작업 ○ 축산물의 출하를 위한 선별, 포장작업	
	구입및판매 ○ 농자재의 구입 및 농축산물의 판매		
	기획관리 ○ 영농설계, 영농일지 작성, 영농교육참석 등		
	기타간접노동 ○ 위에서 분류되지 아니한 일체의 간접노동		

2. 사후 패널화

가. 농가경제조사 자료의 특성

농가경제조사는 5년마다 실시하는 농업총조사를 모집단으로 해서 표본설계를 한 후에 지방통계청에서 현장조사를 실시하여 자료수집하고 있다. 5년간 동일한 표본을 사용하여 매년 집계한 후 ‘농가경제통계’ 보고서와 원시자료로 제공하여 활발히 이용되고 있다.

동 조사는 기본적으로는 농가가 이사를 가거나 농업을 그만 두지 않는 한 5년간(표본교체 주기)은 동일 농가를 조사한다는 점에서 패널화 하였을 때 웨이브간 표본손실이 가장 적어 패널과 비슷한 형태로 다양한 분석 및 활용이 가능할 것으로 보인다.

또한 이러한 조사는 더 다양한 분석이 가능한 패널조사의 성격을 포함하고 있기 때문에 조사된 자료를 활용하여 패널데이터화 한다면, 새로운 패널조사를 설계하는데 드는 많은 비용과 시간을 절감할 수 있을 것이다. 이러한 인식하에 2003년부터 2007년까지 5개년의 자료를 사후 패널화하는 시도해 보고자 한다.

나. 패널구축 방법

1) 가구명부 작성

농가경제조사 최근 5개년간(2003~2007년)의 자료 3,200가구를 대상으로 패널화를 시도하였다. 먼저 농가의 기본정보가 들어있는 nonga파일에서 시도구분코드가 불분명함을 나타내는 ‘00’을 제거한 후 농가연계키(hid)를 생성하였다. hid는 시도(sido-2자리), 지구(jigu-3자리), 농가(nonga-3자리)를 이용하여 생성하되, 패널화 대상 자료 중 가장 최근 자료인 2007년 자료에서 농가(nonga)의 3자리 중 세 번째 자리수를 제외한 8자리 별도의 키를 만들어 가구번호의 중복횟수를 체크한다. -농가(nonga)는 3자리로 구성되지만 마지막 세 번째 자리의 경우 농가의 교체

횃수를 나타내고 있어, 앞의 2자리만을 이용했다.- 이는 교체되었음에도 여전히 존재하고 있는 기존 가구를 제거하기 위한 작업이다. 중복 가구의 경우, 최종 교체가구를 제외한 가구 데이터를 삭제하면 2,589의 데이터가 남게 되는데, 이는 2003년부터 2007년까지 5개년간 바뀌지 않은 표본가구수를 의미한다.

2,589의 가구에 교체횃수 만큼의 가구번호를 새로 생성하면 4,070의 가구가 최종 확인되고, 이렇게 만들어진 4,070의 가구 hid를 이용하여 2003년부터 2007년까지의 데이터 매칭을 시도한다. 다음으로 통계청 지역통계과에서 별도로 받은 가구주 표본대체사유 데이터를 클리닝 한 후³⁶⁾, 만들어진 농가명부 파일에 매칭하게 된다.

가구명부를 작성한 파일은 다음의 <표 4-6>와 같은 형태로 작성되었다.

<표 4-6> 가구명부 파일의 구조

key			연도별 집계 및 교체 사유					전검업 여부					연도별 가중치					5년간 패널 여부	연도별 시도별 농기본 자료
시도 (2)	지구 (3)	농가(3)	03	04	05	06	07	03	04	05	06	07	03	04	05	06	07		
11	010	0 1 0	0	0	0	0												1	
11	010	0 1 1					1	0										1	
11	010	0 2 0	0	0	0	0	0	0										0	
11	010	0 3 0	0																
11	010	0 3 1		2	0	0	0												

- 주 : 1) 농가번호의 셋째자리가 0이면 교체가 없었음을 뜻하고, 1이면 1번 교체, 2이면 2번 교체 등을 의미함.
 2) 연도별 집계에서 Y는 집계, N은 미집계를 의미함.
 3) 교체사유는 1 단독, 2 불능, 3 불응, 4 비농가, 5 장기출타, 6 전출, 7 이상치, 8 기타를 의미함.
 4) 통계청 지역통계과의 표본대체 자료 활용하였음.

36) 지역통계과의 자료는 월별로 정리되어 있고 별도의 코드를 사용하지 않고 있으므로 클리닝 작업이 필요하다.

2) 가구원 명부 작성

현재 가구원의 표본이탈 사유 및 전입 사유는 조사되지 않고 있어 인구통계학적 가구원 분석은 한계가 있다. 또한 가구원의 경우 적합한 연계키를 만들기가 용이하지 않은 문제점을 내포하고 있다. 농가의 경우 명확한 연계키를 만들 수 있었지만, 가구원은 연계키의 역할을 할 대표 변수가 존재하지 않는다. 연계키를 찾기 위한 여러 시도 중 hid, 생년, 성별을 나열한 키가 가장 활용성이 높은 편이나 이 또한 완벽한 일치를 기대하기 어렵다. 생년의 경우, 호적 생년이 아닌 실제 생년을 기재하게 되어 있어 동일 가구원임에도 매칭이 안되는 경우가 일부 발생하기 때문이다.

이러한 문제점에도 hid, 생년, 성별을 연계키로 이용하고자 한다면, 별도의 클리닝 작업이 필요하고, 이 경우에도 데이터 이용에 있어서 한계가 있음을 인식해야 한다. 다음의 <표 4-7>은 가구원명부 파일의 구조를 보여주고 있다.

<표 4-7> 가구원 명부파일의 구조

key			1번가구원					2번가구원					비고
시	지	농	03	04	05	06	07	03	04	05	06	07	
도	구	가	여	여	여	여	여	여	여	여	여	여	여
(2)	(3)	(3)	부	부	부	부	부	부	부	부	부	부	부

- 주 : 1) 가구원 명부 작성 방법은 농가번호 셋째 자리 0인 가구를 기준으로 하여, 생년월, 성별 기준으로 일치화를 점검함.
 2) 사유는 1 전입, 2 전출, 3 사망을 의미함.
 3) 가구원 현황에 대한 자료는 경영주와의 관계, 가구주, 성별, 교육정도, 농업종사 여부, 노동능력, 주업상황(산업, 직업) 취식상태, 전입월, 전출월 등임.

다. 패널화 작업 결과

<표 4-8>는 농가가구의 교체현황을 보여주고 있다. 5개년 동안 교체되지 않은 경우는 매년도 조사된 농가 3,200가구의 80.9%인 2,589가구였으며, 이 중에서 이상치인 47가구를 제외하면 최종적으로 5개년도 지속가구수는 2,542가구로 나타났다.

<표 4-8> 5년간(2003~2005) 패널화 농가교체 현황(단위 : 가구)

	전체	교체 없음	교체 있음	1번	2번	3번	4번	5번	패널	이상 치
가 구 수	3,200	2,589	611	412	149	42	6	2	2,542	47
교체비율 (교체있음비율)	100.0	80.9	19.1 (100.0)	12.9 (67.4)	4.7 (24.4)	1.3 (6.9)	0.2 (1.0)	0.1 (0.3)	79.4	1.5
교체가구수			870	412	298	126	24	10		

<표 4-9>는 교체사유별로 교체가구를 구분하여 제시하고 있다. 5년 합계를 기준으로 가장 많은 교체사유는 단독가구로의 전환이었으며 이러한 경우는 전체 변동가구수 870가구의 28.0%에 해당되는 244가구였다. 다음으로 비농가로의 전환이 전체변동 중에 23.2%를 차지하였으며, 불응의 경우는 전체의 12.0%로 나타나 상대적으로 매우 적었다. 불응이 경우 2004년이후에는 지속적으로 동 경우가 감소하고 있는데 이는 조사에 참여하기 시작한 2년차 이후부터는 불응가구가 점차 감소한다는 일반적인 패널조사에서 나타나는 현상과 유사하다고 할 수 있다.

<표 4-9> 교체사유별 교체 현황(단위 : 가구)

구 분	2003	2004	2005	2006	2007	5년 합계		
조사 가구수(A)	3,200	3,200	3,200	3,200	3,200			
변동 가구수(B)	189	184	209	190	98	870	구성비	
변동율(%)	5.9	5.8	6.5	5.9	3.1	5.4	(%)	
변동 사유	단 독	51	59	55	47	32	244	28.0
	불 능	27	4	6	3	0	40	4.6
	불 응	26	32	28	13	5	104	12.0
	비 농 가	50	40	43	46	23	202	23.2
	장기출타	2	1	6	2	0	11	1.3
	전 출	29	24	26	21	9	109	12.5
기 타	4	24	45	58	29	160	18.4	
변동없는 가구수 (A-B)	3,011	3,016	2,991	3,010	3,102			

<표 4-10>은 가구원 명부 패널화 작업의 결과를 보여주고 있다. 전체 가구원 9,034명중에서 5년 동안 변동이 없는 가구원은 6,683명으로 나타나 전체 가구원의 74.0%로 나타났다. 동 비율이 가구의 경우에는 80.9%였음을 감안할 때 가구원의 변동은 가구의 변동보다는 많은 것으로 나타났다.

가구원변동의 사유로는 전출가구원인 경우가 1,491명으로서 변동이 있었던 가구원의 63.4%였으며, 전입가구원인 경우는 597명으로서 25.4%였으며, 전입전출이 5년 동안 일어난 가구원은 263명으로 11.2%로 나타났다.

<표 4-10> 가구원 명부 패널화 작업 결과(단위 : 가구)

구 분	전체	변동없음	변동있음	전출가구원	전입가구원	기타
가 구 원	9,034	6,683	2,351	1,491	597	263
변동비율(%)	100.0	74.0	26.0	16.5	6.6	2.9
(변동있음비율)			(100.0)	(63.4)	(25.4)	(11.2)

주 : 1) 전출가구원에는 전출, 사망이 포함되며, 전입가구원에는 전입, 출생이 포함되어 있음.

2) 기타는 5년 동안 전입전출이 발생한 가구원임

3. 표본유지 및 변동분석

<표 4-11>은 표본유지율을 제시하고 있다. 인접한 2개 연도를 기준으로 계속 표본유지율은 94~96%의 매우 높은 수준을 보여주었다. 예를 들어 2003년 응답가구 중에서 2004년에도 유지된 가구의 비율은 94.9%였으며, 2004년 응답가구 중에서 2005년에도 유지된 가구는 94.6%였다.

한편, 2003년에 응답한 가구는 시간이 계속적으로 응답한 비율은 2004년 94.9%, 2005년 89.9%, 2006년 86.2%, 2007년 83.6%로 나타났다.

<표 4-11> 기본적인 응답률

2003	2004	2005	2006	2007	기준 년도 가구수	가구수	응답률1	응답률2
Y						3,042		
Y	Y				3,042	2,888	94.94	
	Y	Y			3,056	2,891	94.60	
		Y	Y		3,048	2,912	95.54	
			Y	Y	3,074	2,979	96.91	
Y	Y				3,042	2,888	94.94	94.94
Y	Y	Y			2,888	2,736	94.74	89.94
Y	Y	Y	Y		2,736	2,622	95.83	86.19
Y	Y	Y	Y	Y	2,622	2,542	96.95	83.56

주 : 1) Y는 응답하였음을 의미함.

2) 응답률1은 기준년도 가구수 대비 당해 연도 응답률이며, 응답률2는 5년간 계속 응답한 가구의 2003년 대비 응답률을 의미함.

농가경제조사의 패널유지율은 다른 패널조사 또는 다른 가구패널조사와 비교할 때 매우 높은 표본유지율로 평가된다. <표 4-12>에 나타난 각국의 주요 가구패널조사의 표본유지율을 보면 5차 년도를 기준으로 미국 PSID는 81%, 독일 GSEP는 81%, BHPS는 75% 등으로 나타난다. 한국의 가구패널조사의 경우 대우패널의 5년 유지율은 60%, 노동패널의 경우에는 76%로 나타난다. 이와 같은 가구패널조사는 표본을 추적하여

조사한다는 점에서 농가경제조사와 같은 거처패널보다는 더 높은 표본 유지율을 예상할 수 있다. 그럼에도 불구하고 농가경제 패널화 데이터의 5년 표본유지율이 83.6%에 달한다는 점은 매우 놀라운 사실이라고 할 것이다.

<표 4-12> 주요 가구패널조사에서의 표본유지율

구 분	PSID	GSEP	BHPS	KHPS	KLIPS
2차	89	90	88	79	88
3차	86	86	83	68	81
4차	84	85	79	62	77
5차	81	81	75	60	76
6차	79	79	74	-	77

주 : PSID는 미국가구패널, GSEP는 독일가구패널, BHPS는 영국가구패널, KHPS는 대우경제연구소의 가구패널, KLIPS는 한국노동연구원의 노동패널을 의미함.

다음의 <표 4-13>는 통계청의 경제활동인구조사 패널화 자료와 도시가계조사 패널화 자료의 표본유지율을 제시하고 있다. 경제활동인구조사 패널화 자료의 경우 가구 단위 5년 유지율은 46.5%, 개인 단위 5년 유지율은 36.7%를 기록하고 있다. 도시가계조사 패널화 자료의 경우 가구 단위 5년 유지율은 34.5%, 개인단위 5년 유지율은 27.9%를 기록하고 있다. 경제활동인구조사나 도시가계조사 등과 비교할 때 농가경제조사는 매우 높은 표본유지율을 기록하고 있음을 확인할 수 있다.

<표 4-13> 경제활동인구조사와 도시가계조사에서의 표본유지율

구 분		1998년 1월	1999년 1월	2000년 1월	2001년 1월	2002년 1월	2002년 12월	
경향 표본	가구	패널규모	29,271	23,570	19,448	16,902	15,081	13,616
		유지율		80.5	66.4	57.7	51.5	46.5
	개인	패널규모	93,726	72,563	57,125	47,221	39,931	34,386
		유지율		77.4	60.9	50.4	42.6	36.7
도시 가계 표본	가구	패널규모	4,612	3,572	2,764	2,233	1,868	1,590
		유지율		77.5	59.9	48.4	40.5	34.5
	개인	패널규모	17,258	12,769	9,506	7,360	5,906	4,818
		유지율		74.0	55.1	42.6	34.2	27.9

자료 : 이지연(2005), 가구조사 자료의 중단화 방안, 통계청 통계연구보고서 05-03

4. 중단면 가중치 작성

일반적으로 패널자료가 대표성을 가지려면 각 표본마다 적절한 가중치를 부여할 필요가 있다. 농가경제조사의 패널화 자료를 이용하여 전체 모집단의 특성을 유추하려면 적절한 가중치를 부여하여야 한다. 농가경제조사의 가중치를 부여하기 위해서 먼저 표본들의 응답확률을 구하고 응답률의 역수를 이용하여 인접연도의 가중치를 계산한 다음 최종적으로 모집단의 수와 일치시키는 비율조정을 실시하였다.

먼저 응답률을 추정하기 위해 세 가지 로짓모형(Logit Model)을 고려하였다. 세 모형에 공통적으로 사용된 설명변수는 경영주연령, 경영주연령 제곱, 가구원수, 농업상시종사자, 경지면적, 농가소득, 농업소득, 농가자산, 농가부채, 농업소득의 가계비 충족도 등을 사용하였다. 이러한 모형을 기본모형이라고 하고, 이 기본모형에 전·겸업여부를 나타내는 더미변수를 추가한 모형과 전·겸업의 경우를 구분하여 각각 기본모형을 추정하는 모형 등 세 가지 종류의 모형을 고려하였다.

이 때 나타난 결과는 크게 상이하지는 않았으나, 본고에서는 편의상 세 번째 모형, 즉 전·겸업을 구분하고 각각의 경우에 기본모형을 추정하는 세 번째 모형을 사용한 결과를 제시한다³⁷⁾. 세 번째 모형에서 구체적인 응답률 추정 결과는 다음의 <표 4-14>과 같다.

37) 각각의 경우에 총수를 조정하기 전 가중치의 평균과 분산은 다음과 같다. hwXX_1은 기본모형인 경우, hwXX_2는 기본모형에 전·겸업 더미변수를 추가한 경우, hwXX_3는 전·겸업별로 구분하여 기본모형을 적용한 경우를 의미한다.

구 분	관측치수	평 균	표준편차	최소값	최대값
hw04_1	2,888	379.35	232.39	1.09	1,539.16
hw04_2	2,888	379.26	232.16	1.09	1,535.08
hw04_3	2,888	379.09	231.85	1.08	1,382.56
hw05_1	2,736	399.15	247.71	1.12	2,156.37
hw05_2	2,736	399.04	247.49	1.11	2,205.71
hw05_3	2,736	399.21	248.20	1.09	1,822.85
hw06_1	2,622	416.04	260.17	1.13	2,649.68
hw06_2	2,622	415.90	260.24	1.12	2,806.47
hw06_3	2,622	415.81	259.14	1.11	2,034.44
hw07_1	2,542	430.17	273.16	1.15	2,839.85
hw07_2	2,542	429.96	273.23	1.14	3,015.61
hw07_3	2,542	430.05	273.32	1.13	2,513.93

추정결과를 보면 먼저 2003년도에 응답하고 2004년도에도 응답한 전업농가의 경우를 보면 경영주연령변수와 가구의 상시농업종사자수 그리고 농업소득의 가계비 충족도가 많거나 높은 경우에 통계적으로 유의하게 응답확률이 높아지는 것으로 나타났다. 이에 반해 1종 겸업의 경우에는 가구원수, 농가자산 등이 응답확률을 높이는 유의한 변수로 나타났다으며, 2종 겸업의 경우에는 가구원수가 응답확률을 높이는 유의한 설명변수로 나타났다. 전업이나 전·겸업에 따라서 유의한 설명변수가 상이하았으며, 경영주 연령의 경우에는 전업의 경우에만 유의한 설명변수로 나타났다.

2003년과 2004년에 응답하고 2005년도에도 응답한 가구와 2003년과 2004년에 응답하고 2005년에 응답하지 않은 경우를 비교하면 전업농가의 경우에 경영주 연령, 농업상시 종사자수, 경지면적 등이 응답률을 높이는 유의한 설명변수로 나타났다. 농가자산의 경우에는 응답률을 낮추는 유의한 설명변수였다. 1종 겸업의 경우에는 경영주 연령, 가구원수 등이 응답률을 높이는 유의한 설명변수, 농가자산과 농가부채가 응답률을 낮추는 유의한 설명변수로 나타났다. 2004년에 비해 2005년도 응답률을 결정하는 통계적으로 유의한 요인들이 더욱 많았으며, 이는 상대적으로 체계적인 표본탈락이 있었다고 해석할 수 있을 것이다.

2006년도 전업농가의 경우에는 경영주 연령, 농업상시 종사자수는 응답률을 높이는 유의한 요인으로, 농가소득의 경우에는 응답률을 낮추는 유의한 요인으로 나타났다. 1종 겸업의 경우에는 경영주 연령만이 유의한 변수로 나타났고, 2종 겸업의 경우에는 경영주 연령, 농가소득, 농가자산 등이 유의한 변수로 나타났다. 2007년도 전업농가의 경우에는 농업상시 종사자수, 경지면적, 농가소득 등이 유의한 변수로 나타났다. 1종 겸업의 경우에는 농업상시 종사자수, 농가부채가 유의한 변수로 나타났으며, 2종 겸업의 경우에는 농가소득, 농가자산 등이 유의한 설명변수로 나타났다.

이와 같은 응답률 결과를 종합해 보면 모든 연도에 걸쳐서 전반적으로 응답확률을 높이는 변수는 경영주연령으로 나타났다. 전업농가의 경우에는 농업상시 종사자수가 많을수록 응답확률이 높은 것으로 나타났다. 또한 경지면적이거나 농가소득, 농가자산 등이 유의한 설명변수로 나타난

경우가 있는 데, 이 때 경지면적은 응답률을 높이는 방향으로, 농가소득이나 농가자산의 경우에는 응답률을 낮추는 방향으로 작용하였다. 겸업의 경우에는 전업의 경우와는 달리 매년도 지속적으로 응답률에 영향을 미치는 요인은 거의 없는 것으로 나타났다. 다만 2종 겸업의 경우보다 1종 겸업의 경우가 보다 체계적으로 응답률에 영향을 미치는 요인이 많은 것으로 나타났다.

〈표 4-14〉 응답률 추정모형

(1) 2003-> 2004년

cr34	전 업			1종겸업			2종겸업		
	계수	S.E	p-값	계수	S.E.	p-값	계수	S.E.	p-값
mage2003	0.250	0.091	0.006	0.021	0.223	0.926	0.042	0.135	0.753
mage03s	-0.002	0.001	0.004	0.000	0.002	0.928	0.000	0.001	0.725
g012003	-0.031	0.134	0.817	0.583	0.248	0.019	-0.205	0.106	0.053
g022003	0.947	0.243	0.000	0.260	0.390	0.505	0.258	0.213	0.226
g032003s	0.003	0.003	0.349	0.004	0.004	0.273	0.014	0.009	0.128
g052003s	-0.009	0.010	0.354	0.009	0.018	0.632	0.007	0.009	0.421
g082003s	-0.003	0.015	0.848	-0.020	0.034	0.548	-0.001	0.038	0.970
g172003s	-0.001	0.001	0.187	-0.002	0.001	0.007	0.000	0.001	0.675
g182003s	0.000	0.003	0.994	-0.004	0.004	0.347	-0.001	0.004	0.766
g212003	0.004	0.002	0.058	0.005	0.007	0.515	0.000	0.008	0.961
_cons	-5.485	2.775	0.048	-0.759	6.206	0.903	1.501	3.772	0.691
관측치수	1,735			554			753		

(2) 2004->2005년

cr45	전 업			1종겸업			2종겸업		
	계수	S.E.	p-값	계수	S.E.	p-값	계수	S.E.	p-값
mage2004	0.260	0.113	0.021	0.503	0.148	0.001	0.090	0.151	0.549
mage04s	-0.002	0.001	0.031	-0.005	0.001	0.000	-0.001	0.001	0.561
g012003	0.201	0.173	0.245	0.381	0.213	0.074	0.226	0.141	0.108
g022003	0.574	0.286	0.045	-0.334	0.384	0.385	0.487	0.221	0.027
g032003s	0.011	0.005	0.016	0.003	0.003	0.432	0.010	0.010	0.285
g052003s	0.005	0.013	0.715	-0.016	0.013	0.237	-0.008	0.007	0.311
g082003s	0.002	0.020	0.937	0.008	0.023	0.716	0.027	0.037	0.474
g172003s	-0.002	0.001	0.000	-0.001	0.001	0.023	0.001	0.001	0.522
g182003s	0.003	0.003	0.349	-0.008	0.003	0.019	0.006	0.006	0.311
g212003	0.002	0.003	0.461	0.003	0.005	0.631	0.002	0.008	0.785
_cons	-7.479	3.627	0.039	-9.336	4.279	0.029	-1.553	4.348	0.721
관측치수	1,388			740			760		

〈표 4-14〉 계속

(3) 2005->2006년

cr56	전 업			1종겸업			2종겸업		
	계수	S.E.	p-값	계수	S.E.	p-값	계수	S.E.	p-값
mage2005	0.280	0.122	0.022	0.355	0.210	0.092	0.321	0.151	0.034
mage05s	-0.003	0.001	0.006	-0.003	0.002	0.103	-0.003	0.001	0.039
g012003	-0.036	0.157	0.819	0.055	0.266	0.836	0.214	0.171	0.211
g022003	0.899	0.294	0.002	0.748	0.463	0.106	0.385	0.255	0.131
g032003s	0.000	0.002	0.846	-0.001	0.005	0.793	0.002	0.008	0.770
g052003s	-0.023	0.010	0.021	-0.008	0.015	0.597	0.039	0.014	0.004
g082003s	0.009	0.016	0.584	0.020	0.028	0.473	-0.028	0.049	0.562
g172003s	-0.001	0.001	0.307	0.000	0.002	0.837	-0.001	0.001	0.053
g182003s	-0.001	0.004	0.768	0.004	0.007	0.623	0.000	0.005	0.944
g212003	0.002	0.003	0.491	-0.002	0.006	0.740	0.001	0.008	0.937
_cons	-4.841	4.038	0.231	-7.875	6.122	0.198	-8.192	4.421	0.064
관측치수	1,363			664			709		

(4) 2006->2007년

cr67	전 업			1종겸업			2종겸업		
	계수	S.E.	p-값	계수	S.E.	p-값	계수	S.E.	p-값
mage2006	0.167	0.156	0.281	0.321	0.302	0.287	0.201	0.217	0.354
mage06s	-0.001	0.001	0.242	-0.003	0.002	0.207	-0.002	0.002	0.281
g012003	0.079	0.202	0.695	-0.053	0.378	0.889	-0.227	0.174	0.192
g022003	0.572	0.345	0.097	1.616	0.498	0.001	0.539	0.328	0.101
g032003s	0.013	0.006	0.037	0.006	0.009	0.511	-0.004	0.010	0.722
g052003s	-0.022	0.011	0.053	0.018	0.036	0.615	0.033	0.018	0.068
g082003s	0.018	0.021	0.384	-0.007	0.049	0.889	-0.004	0.042	0.929
g172003s	0.001	0.001	0.474	0.006	0.004	0.142	-0.002	0.001	0.022
g182003s	0.000	0.005	0.924	-0.016	0.007	0.024	0.007	0.009	0.453
g212003	-0.001	0.003	0.762	-0.001	0.006	0.917	-0.003	0.008	0.728
_cons	-3.251	5.147	0.528	-7.168	9.466	0.449	-1.645	6.479	0.800
관측치수	1,229			718			675		

주 : 변수명은 다음과 같음. mage는 경영주 연령, mages는 경영주 연령의 제곱, g01은 가구원수, g02는 농업상시종사자수, g03는 경지면적, g05는 농가소득, g08은 농업소득, g17은 농가자산, g18은 농가부채, g21은 농업소득의 가계비 충족률을 각각 의미함. g로 시작되는 변수의 끝에 붙어 있는 s는 각 변수를 1,000으로 나눈 값을 의미함.

이제 이와 같이 추정된 응답확률을 이용하여 종단면가중치를 구하였다. 먼저 2003년의 종단면가중치는 2003년의 횡단면가중치를 그대로 사용하였다. 2004년의 경우에는 2003년에 응답하고 2004년에도 응답한 가구에 대하여 위에서 추정된 응답률의 역수를 곱하여 2004년의 종단면가중치를 계산하였다. 이후 연도의 경우에도 동일하게 가중치를 작성하였다. 이렇게 작성된 가중치를 종단면 기본가중치라고 명명한다.

매년도의 종단면기본 가중치에 대하여 지구별, 전·겸업별 종단면 기본 가중치의 합을 구하고, 이를 지구별, 전·겸업별 횡단면가중치의 합과 일치하도록 비율 조정을 실시하였다³⁸⁾. 이 때 각 셀에서 표본수가 20개 미만인 경우에는 셀을 지구별로 결합하여 하나의 셀을 만든 다음 비율조정을 실시하였다. 비율조정을 실시하였을 때의 비율과 최종적으로 구해진 가중치의 기초 통계량은 다음과 같다. <표 4-15>에 의하면 2004년의 경우 평균적으로는 종단면기본가중치에 0.979를 곱하여 종단면 최종가중치를 계산하였다. 이러한 비율조정 값은 최소값이 0.583이었고, 최대값이 1.533으로 나타났다. 다른 년도의 경우에도 동일하게 해석할 수 있다.

<표 4-16>는 비율조정을 실시한 이후의 종단면 최종가중치(이하 종단면가중치)의 기초통계량을 제시하고 있다. 2004년의 경우 2,888개의 표본가구에 대하여 적용된 가중치는 평균 371.545가구였으며, 최소값은 0.830, 최대값은 1,362.349로 나타났다³⁹⁾. 종단면 가중치의 평균은 년도가 거듭할수록 커지게 되는데 이는 웨이브가 지남에 따라 표본탈락이 발생하는 반면 표본이 모집단으로 사용하고 있는 농가규모는 약 107만 가구로 변동이 거의 없기 때문이다.

38) 전업과 1종 겸업, 2종 겸업 등 세 가지로 구분하여 보기도 하였으나, 이렇게 구분되는 경우 매우 많은 셀에서 20가구 미만이 나타나 전업과 전겸업의 두 가지로 구분하였다.

39) 통상적인 가중치의 변동보다는 매우 큰 변동이라고 할 수 있다. 이러한 가중치의 변동은 통계청에서 사용하는 기본가중치의 변동폭이 동일 년도의 경우 0.61에서 1,319.53에 달한다는 점에 기인하는 부분이 있다고 하겠다.

〈표 4-15〉 비율조정변수의 기초통계량

변 수	관측치수	평 균	표준편차	최소값	최대값
p2004	2,888	0.979	0.241	0.583	1.533
p2005	2,736	0.983	0.214	0.563	1.441
p2006	2,622	0.970	0.249	0.562	1.544
p2007	2,542	0.971	0.190	0.616	1.485

주 : 횡단면가중치/종단면가중치. pXXXX는 해당년도의 비율조정계수

〈표 4-16〉 가중치의 기초 통계량

변 수	관측치수	평 균	표준편차	최소값	최대값	가중치합
w2004	3,056	351.390	245.498	0.610	1,319.529	1,073,849
w2005	3,048	352.312	240.062	0.758	1,443.839	1,073,848
w2006	3,074	345.879	250.606	0.904	2,040.650	1,063,233
w2007	3,121	341.345	236.977	0.690	1,493.722	1,065,336
hw04_f	2,888	371.545	246.970	0.830	1,362.349	1,073,022
hw05_f	2,736	392.706	255.943	0.891	1,344.533	1,074,443
hw06_f	2,622	405.388	274.860	0.795	1,666.593	1,062,928
hw07_f	2,542	419.094	279.627	0.872	2,669.542	1,065,338

주 : wXXXX는 XXXX년도의 횡단면가중치(통계청 작성), hwXX_f는 본고에서 작성한 XX년의 종단면가중치를 의미함.

〈표 4-17〉은 가중치를 적용하여 계산한 기본 변수의 요약통계량을 제시하고 있다. 먼저 가구원수의 경우에는 종단면가중치를 이용하는 경우와 횡단면가중치를 이용하는 경우에 약 1~2% 정도의 차이가 발생하는 것으로 나타났다. 농업상시종사자의 경우에도 1~2%의 차이가 발생하여 상대적으로 작은 차이를 보여주었다. 경지면적의 경우에는 종단면가중치를 사용하는 경우에는 횡단면가중치를 사용하는 경우보다 2~4%가 많은 것으로 나타났다.

농가소득의 경우 횡단면가중치를 사용하는 경우에 대비하여 종단면가중치를 사용하는 경우의 비율은 2004년에는 100.17%로 거의 차이가 없었으며, 2005년에는 99.85%, 2006년에는 99.44%로서 차이가 거의 미미하였으며, 2007년의 경우에는 98.88%로 1.12%가 차이가 있다.

반면 농업소득의 경우에는 반대로 종단면가중치를 사용하는 경우가 5~9%정도 많은 것으로 나타났다. 이러한 차이는 년도가 지남에 따라 점점 커져서 2004년에는 동 비율이 105.80%였으나, 2005년에는 106.85%, 2006년에는 107.71%, 2007년에는 109.05%로 나타났다.

농가자산의 경우에는 종단면가중치를 사용하는 경우가 적게 나타나는데 동 비율은 2006년과 2007년에 각각 94.69%, 94.32%를 나타냈다. 농가부채의 경우에는 양자의 차이가 거의 없어서 약 1~2% 차이만이 있는 것으로 나타났다.

절대수준 면에서 볼 때 두 가지 가중치를 사용한 결과가 다소 차이를 보이지만 양자의 방향성은 모두 일치하였다.⁴⁰⁾

40) 한편, 미국 PSID의 경우 가구소득변수에서 종단면가중치를 사용한 결과와 CPS의 횡단면가중치를 사용한 결과의 평균비율은 0.96~1.03으로 나타난다(Gouskova et al, 2008).

〈표 4-17〉 가중치를 이용한 통계량 계산결과

(1) 횡단면 및 종단면 가중치를 이용하여 산출한 결과

구 분		횡 단 면			종 단 면		
		표본수	평 균	표준편차	표본수	평 균	표준편차
가구 원수	2004년	3,056	2.85	1.24	2,888	2.82	1.20
	2005년	3,048	2.83	1.21	2,736	2.80	1.19
	2006년	3,074	2.77	1.20	2,622	2.73	1.16
	2007년	3,121	2.77	1.18	2,542	2.73	1.14
농업 상시 종사자	2004년	3,056	1.77	0.70	2,888	1.80	0.68
	2005년	3,048	1.79	0.67	2,736	1.81	0.68
	2006년	3,074	1.76	0.66	2,622	1.79	0.65
	2007년	3,121	1.80	0.66	2,542	1.83	0.65
경지 면적	2004년	3,056	4,912	5,622	2,888	5,023	5,579
	2005년	3,048	4,794	5,370	2,736	4,936	5,498
	2006년	3,074	4,861	5,995	2,622	5,036	6,071
	2007년	3,121	4,880	6,208	2,542	5,081	6,321
농가 소득	2004년	3,056	29,000,529	24,361,770	2,888	29,051,043	24,737,159
	2005년	3,048	30,503,015	26,066,342	2,736	30,457,614	26,318,368
	2006년	3,074	32,303,483	28,685,988	2,622	32,122,328	26,846,235
	2007년	3,121	31,967,420	33,651,625	2,542	31,609,829	33,220,872
농업 소득	2004년	3,056	12,050,150	18,758,170	2,888	12,749,177	19,687,909
	2005년	3,048	11,815,293	18,941,257	2,736	12,625,097	20,157,868
	2006년	3,074	12,091,582	21,326,650	2,622	13,023,998	19,941,501
	2007년	3,121	10,405,729	20,270,960	2,542	11,347,693	22,942,787
농가 자산	2004년	3,056	243,665,473	253,798,331	2,888	243,464,970	257,222,059
	2005년	3,048	298,007,924	316,258,746	2,736	293,426,153	302,976,143
	2006년	3,074	356,963,214	419,367,622	2,622	338,019,557	356,212,258
	2007년	3,121	395,981,102	491,766,971	2,542	373,481,382	407,993,325
농가 부채	2004년	3,056	26,891,994	52,232,306	2,888	27,012,675	52,002,770
	2005년	3,048	27,209,874	54,526,518	2,736	27,758,322	54,012,751
	2006년	3,074	28,160,551	57,551,368	2,622	28,454,827	56,320,650
	2007년	3,121	29,945,749	61,678,144	2,542	30,196,492	61,646,456

〈표 4-17〉 계속

(2) 횡단면 및 종단면 가중치를 사용한 경우에 평균 비율 비교

구 분		종단면/횡단면		연간 증가율(%)	
		평 균	표준편차	횡단면	종단면
가구원수	2004년	98.95	96.77		
	2005년	98.94	98.35	-0.70	-0.71
	2006년	98.56	96.67	-2.12	-2.50
	2007년	98.56	96.61	0.00	0.00
농업상시종 사자	2004년	101.69	97.14		
	2005년	101.12	101.49	1.13	0.56
	2006년	101.70	98.48	-1.68	-1.10
	2007년	101.67	98.48	2.27	2.23
경지면적	2004년	102.25	99.23		
	2005년	102.96	102.38	-2.41	-1.73
	2006년	103.61	101.28	1.40	2.04
	2007년	104.12	101.81	0.40	0.89
농가소득	2004년	100.17	101.54		
	2005년	99.85	100.97	5.18	4.84
	2006년	99.44	93.59	5.90	5.47
	2007년	98.88	98.72	-1.04	-1.60
농업소득	2004년	105.80	104.96		
	2005년	106.85	106.42	-1.95	-0.97
	2006년	107.71	93.51	2.34	3.16
	2007년	109.05	113.18	-13.94	-12.87
농가자산	2004년	99.92	101.35		
	2005년	98.46	95.80	22.30	20.52
	2006년	94.69	84.94	19.78	15.20
	2007년	94.32	82.96	10.93	10.49
농가부채	2004년	100.45	99.56		
	2005년	102.02	99.06	1.18	2.76
	2006년	101.04	97.86	3.49	2.51
	2007년	100.84	99.95	6.34	6.12

이러한 차이가 얼마나 큰 차이인가를 알아보기 위해 농가소득과 농업 소득에 대하여 각각 횡단면가중치를 사용하기 전후, 종단면가중치를 사용

하기 전후의 평균값의 변화를 계산해 보았다. <표 4-17>는 분석결과를 제시하고 있다.

먼저 횡단면가중치와 종단면가중치를 모두 사용하지 않은 경우 농가소득의 횡단면평균 대비 종단면 평균의 비율은 1%이내의 차이를 보여주었다. 농업소득의 경우에는 동 비율이 2~3%의 차이를 보여주었다. 다음으로 횡단면가중치와 종단면가중치를 모두 사용한 경우 농업소득의 횡단면 평균 대비 종단면 평균의 비율은 2004년 100.56%, 2005년 102.46%, 2006년 103.08%, 2007년 105.35%로서 점차 커지는 것으로 나타났다.

한편, 농가소득의 경우 횡단면가중치를 사용하면 가중치를 사용하지 않는 경우에 비해 4~6%정도가 적어지는 것으로 나타났다. 종단면가중치를 사용하는 경우에도 종단면가중치를 사용하지 않는 경우에 비해 약 5~6%정도가 작아지는 것으로 나타났다.

농업소득의 경우 횡단면가중치를 사용하기 전 평균대비 횡단면 가중치를 사용한 후의 평균값의 비율은 2004년 94.49%, 2005년 93.08%, 2006년 94.86%, 2007년 89.56%로 나타났다. 양자의 차이가 약 5~7%차이를 보이다가 2007년에는 약 11%로 확대되었다. 종단면가중치를 이용하여 계산한 동 비율은 2004년 99.41%, 2005년 97.07%, 2006년 99.12%, 2007년 92.71%를 보여주었다.

<표 4-18>의 결과와 각주10에서 밝힌 바와 같이 PSID에서 횡단면가중치와 종단면가중치의 차이가 3내지 4%내외의 차이를 보인다는 점을 감안할 때 <표 4-17>에서 나타난 가중치 부여전후의 결과는 받아들인만 결과라고 하겠다.

<표 4-18> 가중치 사용 전후의 변화분석

(1) 가중치를 사용하지 않은 경우

		관측 치수	평 균	표준편차	최소값	최대값	종단면/ 횡단면
횡 단 면	g052004	3,056	30,737,457	25,713,190	-81,272,071	310,503,207	
	g052005	3,048	32,194,535	28,584,011	-114,000,000	346,682,080	
	g052006	3,074	33,629,511	29,584,692	-92,161,380	528,196,931	
	g052007	3,121	33,522,030	40,408,114	-160,000,000	1,220,000,000	
중 단 면	g052004	2,888	30,805,649	25,789,854	-81,272,071	310,503,207	100.22
	g052005	2,736	32,319,198	28,746,194	-114,000,000	346,682,080	100.39
	g052006	2,622	33,721,939	27,972,429	-32,571,285	379,882,195	100.27
	g052007	2,542	33,398,484	35,424,651	-32,550,380	971,341,660	99.63
횡 단 면	g082004	3,056	12,753,296	19,719,064	-122,000,000	302,780,427	
	g082005	3,048	12,693,991	21,095,436	-147,000,000	206,873,510	
	g082006	3,074	12,746,295	21,734,429	-53,733,739	542,083,281	
	g082007	3,121	11,619,002	24,115,419	-158,000,000	779,317,430	
중 단 면	g082004	2,888	12,824,867	19,644,738	-122,000,000	302,780,427	100.56
	g082005	2,736	13,006,440	21,381,746	-147,000,000	206,873,510	102.46
	g082006	2,622	13,139,169	19,731,059	-53,733,739	245,760,470	103.08
	g082007	2,542	12,240,451	25,037,081	-95,028,012	779,317,430	105.35

(2) 가중치를 사용하여 계산한 경우

		가중치	관측 치수	평 균	표준편차	최소값	최대값	종단면/ 횡단면
횡 단 면	g052004	3,056	29,000,529	24,361,770	-81,272,071	310,503,207		
	g052005	3,048	30,503,015	26,066,342	-114,000,000	346,682,080		
	g052006	3,074	32,303,483	28,685,988	-92,161,380	528,196,931		
	g052007	3,121	31,967,420	33,651,625	-160,000,000	1,220,000,000		
중 단 면	g052004	2,888	29,051,043	24,737,159	-81,272,071	310,503,207	100.17	
	g052005	2,736	30,457,614	26,318,368	-114,000,000	346,682,080	99.85	
	g052006	2,622	32,122,328	26,846,235	-32,571,285	379,882,195	99.44	
	g052007	2,542	31,609,829	33,220,872	-32,550,380	971,341,660	98.88	
횡 단 면	g082004	3,056	12,050,150	18,758,170	-122,000,000	302,780,427		
	g082005	3,048	11,815,293	18,941,257	-147,000,000	206,873,510		
	g082006	3,074	12,091,582	21,326,650	-53,733,739	542,083,281		
	g082007	3,121	10,405,729	20,270,960	-158,000,000	779,317,430		
중 단 면	g082004	2,888	12,749,177	19,687,909	-122,000,000	302,780,427	105.80	
	g082005	2,736	12,625,097	20,157,868	-147,000,000	206,873,510	106.85	
	g082006	2,622	13,023,998	19,941,501	-53,733,739	245,760,470	107.71	
	g082007	2,542	11,347,693	22,942,787	-95,028,012	779,317,430	109.05	

〈표 4-18〉 계속
 (3) 가중치 사용 전후의 평균값과 표준편차의 변화

가중치후/가중치전		평 균	표준편차
횡단면	g052004	94.35	94.74
	g052005	94.75	91.19
	g052006	96.06	96.96
	g052007	95.36	83.28
종단면	g052004	94.30	95.92
	g052005	94.24	91.55
	g052006	95.26	95.97
	g052007	94.64	93.78
횡단면	g082004	94.49	95.13
	g082005	93.08	89.79
	g082006	94.86	98.12
	g082007	89.56	84.06
종단면	g082004	99.41	100.22
	g082005	97.07	94.28
	g082006	99.12	101.07
	g082007	92.71	91.64

주 : g05XXXX, g08XXXX는 각각 XXXX년의 농가소득과 농업소득을 의미함.

제 5 장 농가경제조사 패널화 자료를 이용한 분석결과

1. 소득이동 분석

농가경제조사의 패널화 데이터를 통해 농가 소득의 5분위별 이동경로를 분석해 본다.⁴¹⁾ 먼저 다음의 <표 5-1>은 농가소득과 농업소득을 5분위로 분석한 결과를 제시하고 있다. 농가소득의 5분위 분배율(상위 20%소득/하위20%소득)은 2003년 14.99, 2004년 8.67, 2005년 9.11, 2006년 8.74, 2007년에는 11.36을 보이고 있다. 농가소득의 5분위 분배율은 어떤 추세를 보인다고 보다 매년 급변동하고 있는 특성을 보이고 있으며, 이는 농가소득이 다른 가계와는 달리 매우 변동성이 크다는 특성을 보인다는 점을 보여주는 결과라고 하겠다. 한편, 농업소득의 경우에는 전 기간에 걸쳐서 하위 20%의 평균소득은 마이너스로 나타났다. 특히 2003년과 2007년에는 1분위와 2분위의 평균소득도 마이너스를 기록하고 있다.

<표 5-1> 5분위별 농가소득 및 농업소득 기초 통계량

(1) 농가소득

(단위 :원)

구 분	2003년	2004년	2005년	2006년	2007년
1분위	4,031,518	7,616,246	7,749,792	8,420,918	6,587,579
2분위	14,888,666	16,300,685	16,023,949	17,278,270	16,009,466
3분위	22,048,589	24,093,511	24,589,645	25,575,226	23,938,090
4분위	31,728,608	35,490,527	36,514,410	38,136,080	36,691,214
5분위	60,417,332	66,014,703	70,611,252	73,583,710	74,862,702
전체	26,607,642	29,895,234	31,052,024	32,590,831	31,609,944

41) 이 장에서의 분석은 5개년간 모두 응답한 가구에 대하여 분석한 결과이다.

〈표 5-1〉 계속

(2) 농업소득

(단위 : 원)

구 분	2003년	2004년	2005년	2006년	2007년
1분위	-5,146,358	-1,579,275	-1,901,669	-1,667,724	-3,043,233
2분위	2,935,817	3,763,204	3,128,138	3,145,683	2,343,461
3분위	6,730,414	7,677,037	6,800,890	6,942,683	5,878,327
4분위	13,300,608	14,255,177	13,809,998	14,084,033	12,260,409
5분위	35,402,581	38,741,253	40,558,423	41,291,037	39,362,419
전체	10,631,669	12,552,178	12,469,645	12,745,717	11,347,199

농가소득 기준으로 각 분위별 이동 상황을 제시한 결과가 <표 5-2>이다. 동 표에 의하면 2003년에 1분위이면서 연도가 지남에 따라 계속적으로 1분위에 속하는 가구의 비율은 2004년에 55.8%, 2005년에 49.9%, 2006년 48.8%, 2007년 46.3%로 나타났다. 2003년에 1분위에 속했던 가구 중 약 두 가구 중 한 가구가 5년 후에도 동일하게 1분위에 속하고 있었다. 2003년에 1분위에 속하는 가구 중 최상위 분위인 5분위로 이동한 경우는 2004년 4.6%, 2005년 7.5%, 2006년 7.9%, 2007년 7.0% 등으로 나타났다. 2007년을 기준으로 대략 14,978가구가 5년 만에 소득 최하위 분위에서 소득 최상위 분위로 이동한 것으로 추계된다.

2003년에 5분위이면서 연도가 지남에 따라 동일 분위기를 유지한 가구의 비율은 2004년 54.9%, 2005년 52.9%, 2006년 49.8%, 2007년 47.6%로 나타났다. 결국 5년 후에도 동일한 분위기를 유지한 비율은 1분위와 5분위 모두 약 50%정도인 것으로 추계된다. 2003년에 5분위에 속하는 가구가 2007년에 1분위에 속하는 가구가 된 비율은 7.3%, 15,560가구로 추정되었다. 이와 같은 수치는 1분위에서 5분위로 이동한 경우와 매우 유사하다.

2003년에 3분위이면서 연도가 지남에 따라 동일 분위기를 유지한 가구의 비율은 2004년 36.6%, 2005년 32.5%, 2006년 29.4%, 2007년 30.0%로 나타나 비율이 점차 하락하는 것으로 나타났다. 3분위의 경우에는 1분위와 5분위에 비하여 분위 이동이 많았다. 2003년에 3분위이면서 2분위 이하로 떨어진 비율과 4분위 이상으로 올라간 비율은 2004년에는 각각 29.5%,

34.0%, 2005년에는 34.8%, 33.0%, 2006년에는 37.3%, 33.2%, 2007년에는 35.8%, 34.2%였다. 3분위의 경우 2004년에는 소득분위가 하락하는 비율보다 상향되는 비율이 더욱 높았으나, 2005년의 경우에는 하락하는 비율이 높아지기 시작하였으며, 2006년의 경우에는 양자의 차이가 4.1%p로 가장 커졌다. 2007년의 경우에는 동 차이가 1.6%p로 축소되었으나, 여전히 상향되는 비율보다는 하향되는 비율이 더욱 컸다.

이상과 같은 분석결과는 1분위와 5분위에 속하는 가구의 경우 소득 이동이 상대적으로 매우 적어서 소득계층이 고착화되는 경향이 있음을 시사하고 있다. 이에 반해 3분위에 속하는 가구의 경우 상위 분위로 이동하는 경우보다 하위 분위로 이동하는 경우가 점차 크게 나타났다 (2007년 제외).

농가소득 뿐만 아니라 농업소득의 경우에도 유사한 패턴이 나타났다. 2003년에 농업소득이 1분위에 속하는 가구가 계속 동일 분위에 남아있는 경우는 2004년 59.1%, 2005년 58.1%, 2006년 53.2%, 2007년 48.5%로 나타나 연도가 지남에 따라 비율이 점차 감소하는 것으로 나타났다. 또한 동일 분위에 남아있는 비율은 농가소득의 경우보다는 더욱 큰 것으로 나타나 농가소득에 비해 상대적으로 농업소득의 고착화현상이 나타남을 알 수 있었다. 2003년에 농업소득이 1분위에 속하는 가구가 5분위로 이동하는 비중은 2004년 7.2%, 2005년 9.4%, 2006년 9.3%, 2007년 10.0%로 나타나 연도가 거듭될수록 소폭 비중이 증가하는 것으로 나타났다. 이와 같은 현상은 농업소득이 연도별 가변성이 매우 크다는 점을 보여주는 결과라고 하겠다.

2003년에 농업소득이 5분위 속하는 가구의 경우 동일 분위에 남아있는 비중은 2004년 63.9%, 2005년 64.4%, 2006년 61.6%, 2007년 60.3%로 나타나 연도에 따라서 대개 소폭 하락하는 것으로 나타났다. 그러나 2003년도와 비교하여 2007년에도 동일분위에 남아 있는 비율은 1분위의 경우 48.5%임에 반해 5분위의 경우 60.3%로서 상대적으로 고소득계층일수록 동일 분위에 남아 있는 비율이 컸다. 한편 2003년에 5분위에 있던 가구가 2007년에 1분위로 내려가는 비중은 7.6%로서 1분위에서 5분위로 이동하는 가구 비율 10.0%보다는 작은 것으로 나타났다.

2003년에 농업소득이 3분위에 속하는 가구의 경우 동일 분위에 남아

있는 비중은 2004년 51.3%, 2005년 45.4%, 2006년 42.6%, 2007년 39.1%로 나타나 년도가 지남에 따라 서서히 하락하였다. 다만 이러한 비중은 농가소득에 비해 상대적으로 높은 비중이었다. 2003년에 농업소득이 3분위에 속하는 가구의 경우 1, 2분위로 떨어지는 비중은 2004년 31.47%, 2005년 36.3%, 2006년 31.6%, 2007년 36.5%로 나타나 매년 4분위 이상으로 올라가는 비중에 비해 훨씬 높았다.

이와 같은 분석결과를 종합하면 농가소득의 경우 1분위와 5분위는 동일 분위에 머무는 비중이 상대적으로 높은 것으로 나타나 하위 20%와 상위 20%에서의 소득 고착화현상이 발견되었다. 동일 분위에 머무는 비율은 농가소득보다는 농업소득에서 상위 20%에서 더욱 큰 것으로 나타났다. 반면에 농가소득의 경우 2003년에 1분위에 있다가 2007년에 5분위로 상향 이동한 비중이 7.0%였지만, 농업소득의 경우 동 비중이 10.0%에 달해 상대적으로 농업소득에서 극단적인 계층이동 현상도 많은 것으로 나타났다.

한편, 3분위에 속하는 가구의 경우에는 비교적 소득이동이 활발하게 일어나고 있는데, 농가소득이나 농업소득 모두에서 상향 이동의 비중보다는 하향이동의 비중이 큰 것으로 나타났다.

<표 5-2> 농가소득이동 분석

(1) 2003->2004년

(단위 : 가구, %)

03->04	1분위	2분위	3분위	4분위	5분위	전체
1분위	118,858	46,156	23,550	14,913	9,712	213,189
	55.8	21.7	11.1	7.0	4.6	100.0
2분위	47,306	96,229	42,355	18,428	9,070	213,388
	22.2	45.1	19.9	8.6	4.3	100.0
3분위	24,038	38,754	77,775	48,044	24,164	212,775
	11.3	18.2	36.6	22.6	11.4	100.0
4분위	12,255	19,802	48,631	79,230	53,342	213,260
	5.8	9.3	22.8	37.2	25.0	100.0
5분위	10,844	12,017	20,989	52,121	116,738	212,709
	5.1	5.7	9.9	24.5	54.9	100.0
전체	213,301	212,958	213,300	212,736	213,026	1,065,321
	20.0	20.0	20.0	20.0	20.0	100.0

〈표 5-2〉 계속
(2) 2003->2005년

(단위 : 가구, %)

03->05	1분위	2분위	3분위	4분위	5분위	전체
1분위	106,456	53,108	22,025	15,608	15,992	213,189
	49.9	24.9	10.3	7.3	7.5	100.0
2분위	60,576	75,131	37,748	27,509	12,424	213,388
	28.4	35.2	17.7	12.9	5.8	100.0
3분위	22,559	50,936	69,103	49,617	20,560	212,775
	10.6	23.9	32.5	23.3	9.7	100.0
4분위	11,867	19,373	59,726	71,861	50,433	213,260
	5.6	9.1	28.0	33.7	23.7	100.0
5분위	12,002	14,687	24,290	49,129	112,601	212,709
	5.6	6.9	11.4	23.1	52.9	100.0
전 체	213,460	213,235	212,892	213,724	212,010	1,065,321
	20.0	20.0	20.0	20.1	19.9	100.0

(3) 2003->2006년

(단위 : 가구, %)

03->06	1분위	2분위	3분위	4분위	5분위	전체
1분위	104,109	46,214	25,859	20,221	16,786	213,189
	48.8	21.7	12.1	9.5	7.9	100.0
2분위	52,991	80,289	46,009	23,270	10,829	213,388
	24.8	37.6	21.6	10.9	5.1	100.0
3분위	30,528	48,917	62,644	49,769	20,917	212,775
	14.4	23.0	29.4	23.4	9.8	100.0
4분위	15,755	21,736	52,044	65,198	58,527	213,260
	7.4	10.2	24.4	30.6	27.4	100.0
5분위	9,746	16,287	26,162	54,595	105,919	212,709
	4.6	7.7	12.3	25.7	49.8	100.0
전 체	213,129	213,443	212,718	213,053	212,978	1,065,321
	20.0	20.0	20.0	20.0	20.0	100.0

〈표 5-2〉 계속
(4) 2003년->2007년

(단위 : 가구, %)

03->07	1분위	2분위	3분위	4분위	5분위	전체
1분위	98,785	53,689	23,372	22,365	14,978	213,189
	46.3	25.2	11.0	10.5	7.0	100.0
2분위	52,911	72,822	45,130	28,335	14,190	213,388
	24.8	34.1	21.2	13.3	6.7	100.0
3분위	27,858	48,196	63,920	46,622	26,179	212,775
	13.1	22.7	30.0	21.9	12.3	100.0
4분위	18,003	23,400	47,389	67,969	56,499	213,260
	8.4	11.0	22.2	31.9	26.5	100.0
5분위	15,560	15,386	32,940	47,655	101,168	212,709
	7.3	7.2	15.5	22.4	47.6	100.0
전체	213,117	213,493	212,751	212,946	213,014	1,065,321
	20.0	20.0	20.0	20.0	20.0	100.0

〈표 5-3〉 농업소득이동 분석

(1) 2003->2004년

(단위 : 가구, %)

03->04	1분위	2분위	3분위	4분위	5분위	전체
1분위	125,902	35,598	19,773	16,426	15,369	213,068
	59.1	16.7	9.3	7.7	7.2	100.0
2분위	43,603	109,407	39,133	15,826	5,268	213,237
	20.5	51.3	18.4	7.4	2.5	100.0
3분위	19,108	47,811	94,236	39,283	12,695	213,133
	9.0	22.4	44.2	18.4	6.0	100.0
4분위	10,449	12,645	50,776	96,184	43,284	213,338
	4.9	5.9	23.8	45.1	20.3	100.0
5분위	14,484	7,153	9,116	46,073	135,719	212,545
	6.8	3.4	4.3	21.7	63.9	100.0
전 체	213,546	212,614	213,034	213,792	212,335	1,065,321
	20.1	20.0	20.0	20.1	19.9	100.0

(2) 2003->2005년

(단위 : 가구, %)

03->05	1분위	2분위	3분위	4분위	5분위	전체
1분위	123,740	33,717	19,992	15,691	19,928	213,068
	58.1	15.8	9.4	7.4	9.4	100.0
2분위	51,723	96,750	40,816	17,754	6,194	213,237
	24.3	45.4	19.1	8.3	2.9	100.0
3분위	17,577	59,709	83,238	42,965	9,644	213,133
	8.3	28.0	39.1	20.2	4.5	100.0
4분위	11,623	16,999	56,627	87,799	40,290	213,338
	5.5	8.0	26.5	41.2	18.9	100.0
5분위	8,666	6,316	12,031	48,671	136,861	212,545
	4.1	3.0	5.7	22.9	64.4	100.0
전 체	213,329	213,491	212,704	212,880	212,917	1,065,321
	20.0	20.0	20.0	20.0	20.0	100.0

<표 5-3> 계속

(3) 2003->2006년

(단위 : 가구, %)

03->06	1분위	2분위	3분위	4분위	5분위	전체
1분위	113,318	37,390	27,445	15,214	19,701	213,068
	53.2	17.6	12.9	7.1	9.3	100.0
2분위	50,928	103,412	37,808	14,269	6,820	213,237
	23.9	48.5	17.7	6.7	3.2	100.0
3분위	19,565	47,694	90,735	41,014	14,125	213,133
	9.2	22.4	42.6	19.2	6.6	100.0
4분위	14,394	20,759	43,524	93,319	41,342	213,338
	6.8	9.7	20.4	43.7	19.4	100.0
5분위	15,635	3,933	12,960	49,123	130,894	212,545
	7.4	1.9	6.1	23.1	61.6	100.0
전체	213,840	213,188	212,472	212,939	212,882	1,065,321
	20.1	20.0	19.9	20.0	20.0	100.0

(4) 2003->2007년

(단위 : 가구, %)

03->07	1분위	2분위	3분위	4분위	5분위	전체
1분위	103,385	50,023	20,825	17,446	21,389	213,068
	48.5	23.5	9.8	8.2	10.0	100.0
2분위	48,489	93,662	46,192	17,566	7,328	213,237
	22.7	43.9	21.7	8.2	3.4	100.0
3분위	28,755	49,109	83,314	37,135	14,820	213,133
	13.5	23.0	39.1	17.4	7.0	100.0
4분위	17,164	15,919	50,300	88,719	41,236	213,338
	8.1	7.5	23.6	41.6	19.3	100.0
5분위	16,060	4,169	12,026	52,230	128,060	212,545
	7.6	2.0	5.7	24.6	60.3	100.0
전체	213,853	212,882	212,657	213,096	212,833	1,065,321
	20.1	20.0	20.0	20.0	20.0	100.0

2. 자산, 부채이동 분석

이번에는 자산과 부채관련 변수의 이동을 분석한다. <표 5-4>는 자산과 부채관련 변수의 기초 통계량을 제시하고 있다. 농가자산의 5분위 분배율은 2003년부터 각각 9.14, 9.35, 10.17, 10.75, 11.14 등으로 지속적으로 증가하고 있다. 자산이 가장 많은 5분위의 평균자산액은 2003년에 5억 382만원이었으나, 2007년에는 9억 8032만원에 달하여 1.82배 증가하였다. 동 기간 중 5분위의 농가소득이 1.24배 증가하였음을 감안할 때 자산증가율이 소득증가율에 비해 월등히 높았음을 알 수 있다.

농가부채의 경우 부채가 가장 많은 5분위의 경우에 2003년에는 평균 9,792만원이었으나 2007년에는 1억 1,455만원으로 증가하였다. 이외의 계층에서는 분위별로 평균부채가 큰 변화를 보이지는 않았다.

<표 5-4> 5분위별 농가자산 및 농가부채의 기초통계량

(1) 농가자산

(단위 : 원)

구 분	2003년	2004년	2005년	2006년	2007년
1분위	55,141,537	65,838,695	75,857,228	82,643,845	87,993,231
2분위	100,070,878	120,378,580	141,504,414	153,330,542	163,937,811
3분위	148,356,649	172,894,599	206,445,912	229,080,369	249,279,384
4분위	218,996,721	259,267,954	315,061,279	357,746,772	386,715,597
5분위	503,815,164	615,578,765	771,613,995	888,081,701	980,320,708
전 체	205,208,693	246,603,052	301,938,041	342,019,316	373,479,010

(2) 농가부채

(단위 : 원)

구 분	2003년	2004년	2005년	2006년	2007년
1분위	0	0	0	0	0
2분위	1,263,637	1,215,922	1,185,972	828,049	1,107,435
3분위	8,210,157	8,325,297	8,433,172	7,316,428	8,433,577
4분위	25,930,743	25,684,043	26,427,099	26,337,636	27,490,707
5분위	97,916,892	102,377,376	108,451,727	110,933,398	114,548,470
전 체	26,493,413	27,357,331	28,732,546	28,990,720	30,195,681

<표 5-5>는 자산분위의 이동 상황을 제시하고 있다. 2003년에 자산 1분위에 속했던 농가 중에서 연도가 지남에 따라 동일 분위에 속하는 비중은 2004년 82.5%, 2005년 77.2%, 2006년 73.6%, 2007년 72.5%로 나타나, 동일 분위에 속하는 비중이 농가소득이나 농업소득의 경우에 비하여 매우 높은 것으로 나타났다. 2003년에 자산 1분위에 속했던 농가 중에서 2007년에 자산 5분위에 속하는 농가의 비중은 0.4%에 불과하였다.

2003년에 자산 5분위에 속했던 농가 중에서 연도가 지남에 따라 동일 분위에 속하는 비중은 2004년 85.7%, 2005년 81.1%, 2006년 77.8%, 2007년 78.6%로 나타났다. 자산 1분위의 경우보다 비율이 높았으며, 물론 농가소득이나 농업소득에서의 5분위 경우보다 높게 나타났다. 2003년에 자산 3분위에 속했던 농가 중에서 연도가 지남에 따라 동일 분위에 속하는 비중은 2004년 65.9%, 2005년 56.4%, 2006년 50.1%, 2007년 47.2%로 나타났다. 역시 농업소득이나 농가소득에 비해 상대적으로 동일 분위에 남아있는 비율이 높게 나타났다.

이와 같은 점을 종합하면 예상한 바와 같이 농가소득이나 농업소득보다 농가자산의 경우에는 분위별 이동이 상대적으로 적고, 특히 1분위와 5분위에서 동일 분위에 남아 있는 비율이 매우 크게 나타났다.

<표 5-6>는 5분위별 부채이동 상황을 보여주고 있다. 2003년에 부채가 많은 5분위에 속했던 농가 중에서 2004년에 5분위에 머무는 비중은 86.2%, 4분위로 내려간 비중은 10.7%를 기록하였다. 2005년에는 5분위에 머무는 비중이 80.5%로 하락한 반면 4분위로 내려간 비중은 15.3%, 3분위로 내려간 비중은 2.1%를 기록하였다. 2006년에는 5분위에 머무는 비중이 75.7%로 하락한 반면 4분위로 내려간 비중은 18.4%, 3분위로 내려간 비중은 2.9%를 기록하였다. 2007년에는 5분위에 머무는 비중이 71.3%로 하락한 반면 4분위로 내려간 비중은 20.4%, 3분위로 내려간 비중은 5.6%를 기록하였다.

이와 같은 모습은 부채가 많은 농가의 문제가 서서히 해소되는 모습도 보이고 있지만 해소되는 속도가 매우 완만하게 진행되고 있음을 알 수 있다.

〈표 5-5〉 자산분위의 이동

(1) 2003->2004년

(단위 : 가구, %)

03->04	1분위	2분위	3분위	4분위	5분위	전체
1분위	176,250	32,442	3,488	591	834	213,605
	82.5	15.2	1.6	0.3	0.4	100.0
2분위	33,918	140,965	34,177	3,700	359	213,119
	15.9	66.1	16.0	1.7	0.2	100.0
3분위	2,858	36,595	140,008	26,607	6,461	212,529
	1.3	17.2	65.9	12.5	3.0	100.0
4분위	318	2,751	33,339	154,111	22,536	213,055
	0.2	1.3	15.7	72.3	10.6	100.0
5분위	0	304	1,973	28,296	182,440	213,013
	0.0	0.1	0.9	13.3	85.7	100.0
전 체	213,344	213,057	212,985	213,305	212,630	1,065,321
	20.0	20.0	20.0	20.0	20.0	100.0

(2) 2003->2005년

(단위 : 가구, %)

03->05	1분위	2분위	3분위	4분위	5분위	전체
1분위	164,961	40,610	6,351	945	738	213,605
	77.2	19.0	3.0	0.4	0.4	100.0
2분위	40,541	122,169	39,831	9,743	835	213,119
	19.0	57.3	18.7	4.6	0.4	100.0
3분위	7,048	45,187	119,848	34,767	5,679	212,529
	3.3	21.3	56.4	16.4	2.7	100.0
4분위	109	4,650	44,232	131,083	32,981	213,055
	0.1	2.2	20.8	61.5	15.5	100.0
5분위	468	541	3,313	36,040	172,651	213,013
	0.2	0.3	1.6	16.9	81.1	100.0
전 체	213,127	213,157	213,575	212,578	212,884	1,065,321
	20.0	20.0	20.1	20.0	20.0	100.0

150 농가소득의 동태적 변화요인

<표 5-5> 계속

(3) 2003->2006년

(단위 : 가구, %)

03->06	1분위	2분위	3분위	4분위	5분위	전체
1분위	157,130	47,623	5,964	2,022	866	213,605
	73.6	22.3	2.8	1.0	0.4	100.0
2분위	46,003	105,937	48,981	11,252	946	213,119
	21.6	49.7	23.0	5.3	0.4	100.0
3분위	8,367	50,441	106,539	42,182	5,000	212,529
	3.9	23.7	50.1	19.9	2.4	100.0
4분위	1,408	8,626	48,524	114,015	40,482	213,055
	0.7	4.1	22.8	53.5	19.0	100.0
5분위	468	384	3,137	43,409	165,615	213,013
	0.2	0.2	1.5	20.4	77.8	100.0
전 체	213,376	213,011	213,145	212,880	212,909	1,065,321
	20.0	20.0	20.0	20.0	20.0	100.0

(4) 2003->2007년

(단위 : 가구, %)

03->07	1분위	2분위	3분위	4분위	5분위	전체
1분위	154,905	46,948	8,093	2,260	1,399	213,605
	72.5	22.0	3.8	1.1	0.7	100.0
2분위	47,059	101,954	48,234	14,357	1,515	213,119
	22.1	47.8	22.6	6.7	0.7	100.0
3분위	7,812	55,713	100,287	43,462	5,255	212,529
	3.7	26.2	47.2	20.5	2.5	100.0
4분위	2,929	7,856	53,323	111,639	37,308	213,055
	1.4	3.7	25.0	52.4	17.5	100.0
5분위	468	543	3,147	41,533	167,322	213,013
	0.2	0.3	1.5	19.5	78.6	100.0
전 체	213,173	213,014	213,084	213,251	212,799	1,065,321
	20.0	20.0	20.0	20.0	20.0	100.0

〈표 5-6〉 부채분위의 이동

(1) 2003->2004년

(단위 : 가구, %)

03->04	1분위	2분위	3분위	4분위	5분위	전체
1분위	211,562	29,691	15,479	8,321	1,175	266,228
	79.5	11.2	5.8	3.1	0.4	100.0
2분위	57,285	71,291	32,310	6,668	2,271	169,825
	33.7	42.0	19.0	3.9	1.3	100.0
3분위	19,315	28,421	127,196	24,871	3,530	203,333
	9.5	14.0	62.6	12.2	1.7	100.0
4분위	7,335	2,135	30,643	150,870	22,478	213,461
	3.4	1.0	14.4	70.7	10.5	100.0
5분위	1,753	2,361	2,564	22,619	183,177	212,474
	0.8	1.1	1.2	10.7	86.2	100.0
전체	297,250	133,899	208,192	213,349	212,631	1,065,321
	27.9	12.6	19.5	20.0	20.0	100.0

(2) 2003->2005년

(단위 : 가구, %)

03->05	1분위	2분위	3분위	4분위	5분위	전체
1분위	192,085	34,943	20,763	14,709	3,728	266,228
	72.2	13.1	7.8	5.5	1.4	100.0
2분위	66,742	57,856	34,793	7,594	2,840	169,825
	39.3	34.1	20.5	4.5	1.7	100.0
3분위	29,281	30,658	101,463	33,724	8,207	203,333
	14.4	15.1	49.9	16.6	4.0	100.0
4분위	10,850	4,998	45,826	124,794	26,993	213,461
	5.1	2.3	21.5	58.5	12.7	100.0
5분위	1,500	3,220	4,389	32,406	170,959	212,474
	0.7	1.5	2.1	15.3	80.5	100.0
전체	300,458	131,675	207,234	213,227	212,727	1,065,321
	28.2	12.4	19.5	20.0	20.0	100.0

〈표 5-6〉 계속

(3) 2003->2006년

(단위 : 가구, %)

03->06	1분위	2분위	3분위	4분위	5분위	전체
1분위	177,917	37,753	30,546	12,581	7,431	266,228
	66.8	14.2	11.5	4.7	2.8	100.0
2분위	77,503	37,450	37,043	12,502	5,327	169,825
	45.6	22.1	21.8	7.4	3.1	100.0
3분위	38,654	25,443	86,646	41,050	11,540	203,333
	19.0	12.5	42.6	20.2	5.7	100.0
4분위	21,487	3,598	52,857	107,681	27,838	213,461
	10.1	1.7	24.8	50.5	13.0	100.0
5분위	4,296	2,173	6,052	39,081	160,872	212,474
	2.0	1.0	2.9	18.4	75.7	100.0
전 체	319,857	106,417	213,144	212,895	213,008	1,065,321
	30.0	10.0	20.0	20.0	20.0	100.0

(4) 2003->2007년

(단위 : 가구, %)

03->07	1분위	2분위	3분위	4분위	5분위	전체
1분위	149,961	60,003	33,192	13,782	9,290	266,228
	56.3	22.5	12.5	5.2	3.5	100.0
2분위	66,386	47,548	32,733	17,513	5,645	169,825
	39.1	28.0	19.3	10.3	3.3	100.0
3분위	37,757	33,789	79,037	39,724	13,026	203,333
	18.6	16.6	38.9	19.5	6.4	100.0
4분위	17,870	7,619	56,314	98,478	33,180	213,461
	8.4	3.6	26.4	46.1	15.5	100.0
5분위	4,944	864	11,837	43,370	151,459	212,474
	2.3	0.4	5.6	20.4	71.3	100.0
전 체	276,918	149,823	213,113	212,867	212,600	1,065,321
	26.0	14.1	20.0	20.0	20.0	100.0

3. 전·겸업별 농가이동

다음의 <표 5-7>은 전·겸업별 농가의 이동 상황을 제시하고 있다. 2003년에 전업농가인 경우가 계속 전업농가로 남는 비중은 2004년 72.0%, 2005년 73.4%, 2006년 70.1%, 2007년 72.0%로 나타나 연도별로 큰 변화가 없었다.

2종 겸업의 경우에도 지속적으로 2종 겸업으로 남는 비중이 높게 나타났다. 2003년에 2종 겸업을 하던 농가의 경우 지속적으로 2종 겸업을 하는 비중은 2004년 84.4%, 2005년 77.8%, 2006년 76.4%, 2007년 69.6%로 나타났다. 2종 겸업의 경우 높은 지속 확률을 보이고 있지만 추세적으로 지속 확률이 낮아지고 있음을 알 수 있다.

이에 반해 1종 겸업의 경우에는 상대적으로 이동 비중이 높게 나타났다. 2003년에 1종 겸업을 하던 농가 중에서 지속적으로 1종 겸업을 하는 비중은 2004년 58.1%, 2005년 46.3%, 2006년 46.0%, 2007년 37.2%로 나타났다. 전업이나 2종 겸업 농가에 비해 상대적으로 지속 확률이 매우 낮으며, 동 비율조차도 점차 하락하고 있다.

2003년에 1종 겸업을 하던 농가 중에서 전업농가로 변환한 비율은 2004년 33.5%, 2005년 39.0%, 2006년 40.0%, 2007년 48.0%로 나타나 상당 수의 1종 겸업가구가 전업농가로 전환하고 있으며, 동 비율이 해마다 증가하는 것으로 나타났다. 한편, 전업농가에서 1종 겸업으로 변화한 가구의 비율도 적지 않다. 2003년에 전업농가이다가 1종 겸업 농가로 변화한 비율은 2004년 21.6%, 2005년 19.7%, 2006년 23.0%, 2007년 19.5% 등으로 나타났다. 전업농가와 1종 겸업농가 사이에 상호 이동이 활발하게 일어나고 있음을 알 수 있다.

<표 5-7> 전·겸업별 농가 이동

(1) 2003->2004년

(단위 : 가구, %)

03->04	전 업	1종겸업	2종겸업	전 체
전 업	478,686	143,510	43,101	665,297
	72.0	21.6	6.5	100.0
1종겸업	45,661	79,182	11,551	136,394
	33.5	58.1	8.5	100.0
2종겸업	14,763	26,423	222,444	263,630
	5.6	10.0	84.4	100.0
전 체	539,110	249,115	277,096	1,065,321
	50.6	23.4	26.0	100.0

(2) 2003->2005년

(단위 : 가구, %)

03->05	전 업	1종겸업	2종겸업	전 체
전 업	488,545	131,179	45,573	665,297
	73.4	19.7	6.9	100.0
1종겸업	53,253	63,159	19,982	136,394
	39.0	46.3	14.7	100.0
2종겸업	22,571	35,848	205,211	263,630
	8.6	13.6	77.8	100.0
전 체	564,369	230,186	270,766	1,065,321
	53.0	21.6	25.4	100.0

(3) 2004->2006년

(단위 : 가구, %)

03->06	전업	1종겸업	2종겸업	전체
전업	466,346	152,860	46,091	665,297
	70.1	23.0	6.9	100.0
1종겸업	54,506	62,776	19,112	136,394
	40.0	46.0	14.0	100.0
2종겸업	28,191	34,015	201,424	263,630
	10.7	12.9	76.4	100.0
전체	549,043	249,651	266,627	1,065,321
	51.5	23.4	25.0	100.0

(3) 2003->2007년

(단위 : 가구, %)

03->07	전업	1종겸업	2종겸업	전체
전업	479,148	129,621	56,528	665,297
	72.0	19.5	8.5	100.0
1종겸업	65,527	50,680	20,187	136,394
	48.0	37.2	14.8	100.0
2종겸업	50,591	29,691	183,348	263,630
	19.2	11.3	69.6	100.0
전체	595,266	209,992	260,063	1,065,321
	55.9	19.7	24.4	100.0

4. 농업소득 및 농업순생산 결정요인 분석

농가경제조사의 패널화 자료를 이용하여 농업소득 및 농업순생산의 결정요인을 분석하였다⁴²⁾. <표 5-8>은 분석결과를 제시하고 있다.

42) <부록 2>는 분석에 사용된 변수들의 기초통계량을 제시하고 있다.

<표 5-8>의 (1)은 전 표본과 5년 유지표본의 경우를 구분하여 농업소득의 결정요인을 분석한 결과를 제시하고 있다. 전체 표본을 사용한 임의효과 패널회귀분석에 의하면 분석에 사용된 변수 중에서 닭(c21)변수를 제외하고는 모두 통계적으로 유의하게 농업소득을 결정하는 요인으로 나타났다.

다른 조건이 동일할 때 농업상시종사자수가 1인 증가하는 경우 농업소득은 92만원 정도 증가하는 것으로 나타났다. 농가자산의 회귀계수는 통계적으로 유의한 양수로 나타났으며, 농가부채의 회귀계수는 통계적으로 유의한 음수로 나타났다. 자영노동시간이 길수록 농업소득이 높게 나타났다.

자작지가 많을 수록 농업소득이 많은 것으로 나타났는데, 계수의 상대적 크기는 논면적이 가장 크게 나타났으며, 다음으로 과수원 면적이 컸고, 밭 면적이 상대적으로 가장 작게 나타났다. 차용지 면적도 클수록 농업소득이 높은 것으로 나타났는데, 다만 계수의 크기는 논이나 밭인 경우에는 자작지가 차용지에 비해 크게 나타났으며, 과수원의 경우에는 자작지에 비해 차용지가 높게 나타났다⁴³⁾. 가축류의 경우 한우나 젖소, 돼지는 통계적으로 유의하게 농업소득을 높이는 요인으로 나타났지만, 닭인 경우에는 통계적으로 유의한 요인이 아니었다. 가축류의 경우 마리당 농업소득을 가장 높이는 가축은 젖소인 것으로 나타났다.

경영주의 연령이 많을수록 농업소득도 많아지는 것으로 나타났으며, 다만 농업소득의 증가율은 연령이 많아질수록 낮아지는 것으로 나타났다. 이와 같은 결론은 5년 유지표본을 사용한 경우나 pooled regression을 사용한 경우에도 유사하게 나타났다.

<표 5-8>의 (2)는 전 표본과 5년 유지표본의 경우를 구분하여 농업순생산의 결정요인을 분석한 결과를 제시하고 있다. 전체 표본을 사용한 임의효과 패널회귀분석에 의하면 분석에 사용된 모든 변수가 모두 통계적으로 유의하게 농업순생산을 결정하는 요인으로 나타났다⁴⁴⁾.

농업순생산에서 주목되는 점은 경영주연령의 경우이다. pooled regression에 의하면 경영주 연령 변수는 통계적으로 유의한 변수가 아닌

43) 이 부분은 해석하기 어려운 부분이다. 추가적인 검토가 필요한 것으로 판단된다.

44) 돼지의 경우에는 10%유의수준에서 통계적으로 유의하였다.

것으로 나타났으나, 개인효과를 통제한 이후에는 5%유의 수준에서 통계적으로 유의한 값으로 나타나고 있다. 농업소득과는 달리 농업순생산의 경우에는 전체 표본을 사용하는 경우와 5년 유지 표본을 사용하는 경우에 분석결과가 달리 나타났다. 5년 유지 표본을 사용하여 임의효과 패널회귀분석을 실시한 경우 경영주 연령은 다시 통계적으로 유의하지 않은 변수로 나타났으며, 닭변수(c21)은 통계적으로 유의하지 않게 나타났다. 이러한 결과는 pooled regression에서도 동일하다.

<표 5-8>의 (3)은 농업상시종사자 1인당 농업소득 결정요인을 전체 표본과 5년 유지 표본으로 구분하여 제시하고 있다. 임의효과 패널회귀 분석 모형에 의하면 이 경우에도 닭 변수를 제외하면 분석에 사용된 설명변수가 모두 통계적으로 유의한 변수인 것으로 나타났다. 경영주 연령 효과가 pooled regression결과보다 월등히 높게 나타난 점이 주목된다. 5년 유지 표본의 분석결과와 비교해 보면 5년 유지 표본의 경우에는 가구주 연령변수의 통계적 유의도가 소폭 하락하고 있다. 농업상시종사자 1인당 농업순생산의 경우에도 농업소득의 경우와 거의 유사하다⁴⁵⁾.

45) 다만, 농업상시종사자 1인당 농업순생산 결정식에서 부채의 부호가 플러스로 나온 점은 주목할 만하다. 이 부분에 대해서는 농업경제학적 해석이 필요할 것으로 판단된다.

〈표 5-8〉 회귀분석 결과

(1) 농업소득 - 전체 표본

	pooled			random effect		
	계 수	표본오차	p-값	계 수	표본오차	p-값
g02	1,027,062.000	238,776.200	0.000	915,511.200	283,154.500	0.001
g17	0.005	0.000	0.000	0.006	0.001	0.000
g18	-0.017	0.003	0.000	-0.014	0.003	0.000
g04	4,059.064	118.990	0.000	4,055.144	139.198	0.000
mage	423,493.900	137,430.900	0.002	615,490.400	172,048.000	0.000
mages	-3,908.201	1,157.423	0.001	-5,467.536	1,448.169	0.000
f07	1,443.552	65.278	0.000	1,389.288	81.230	0.000
f08	999.683	103.942	0.000	1,006.452	128.943	0.000
f09	1,623.754	132.737	0.000	1,450.312	167.923	0.000
f12	724.817	47.071	0.000	680.242	58.613	0.000
f13	155.675	54.405	0.004	177.422	66.653	0.008
f14	1,529.076	211.948	0.000	1,707.920	254.037	0.000
c16	280,030.800	14,012.240	0.000	263,497.300	17,996.970	0.000
c17	640,059.800	23,016.530	0.000	607,594.800	29,657.910	0.000
c19	21,817.930	6,087.709	0.000	19,886.850	7,735.144	0.010
c21	187.888	114.673	0.101	71.137	123.850	0.566
_cons	-14,000,000	4,024,644	0.001	-19,500,000	5,037,356	0.000
R ²			0.313			0.312

(1) 농업소득 - 5년 표본

	pooled			random effect		
	계 수	표본오차	p-값	계 수	표본오차	p-값
g02	1,223,423.000	253,918.500	0.000	1,166,176.000	301,501.400	0.000
g17	0.006	0.001	0.000	0.007	0.001	0.000
g18	-0.019	0.003	0.000	-0.014	0.004	0.000
g04	4,195.282	130.555	0.000	4,254.918	154.769	0.000
mage	273,729.500	149,620.200	0.067	448,667.300	188,622.000	0.017
mages	-2,821.073	1,262.218	0.025	-4,228.738	1,589.861	0.008
f07	1,330.146	70.452	0.000	1,220.159	87.289	0.000
f08	854.425	111.822	0.000	851.236	137.034	0.000
f09	1,606.202	135.193	0.000	1,388.513	169.132	0.000
f12	730.140	46.848	0.000	667.756	57.658	0.000
f13	71.494	56.780	0.208	59.051	69.911	0.398
f14	1,405.129	207.617	0.000	1,590.615	247.170	0.000
c16	270,692.400	13,835.910	0.000	249,507.500	17,449.090	0.000
c17	544,583.200	24,359.260	0.000	503,063.700	30,922.930	0.000
c19	23,047.470	5,925.912	0.000	20,685.620	7,404.592	0.005
c21	-4.455	144.284	0.975	-103.995	154.115	0.500
_cons	-9,006,732	4,385,975	0.040	-14,300,000	5,531,502	0.010
R ²			0.328			0.327

(2) 농업순생산 - 전체 표본

	pooled			random effect		
	계 수	표본오차	p-값	계 수	표본오차	p-값
g02	790,462.000	246,412.900	0.001	820,280.800	294,795.800	0.005
g17	0.008	0.000	0.000	0.009	0.001	0.000
g18	0.010	0.003	0.000	0.008	0.003	0.007
g04	7,351.067	122.796	0.000	7,217.085	144.697	0.000
mage	112,961.600	141,826.300	0.426	413,829.900	180,237.600	0.022
mages	-1,829.601	1,194.441	0.126	-4,373.078	1,517.012	0.004
f07	1,650.405	67.365	0.000	1,646.351	85.019	0.000
f08	1,097.463	107.266	0.000	1,107.689	134.844	0.000
f09	2,166.787	136.983	0.000	1,999.037	176.028	0.000
f12	1,668.478	48.577	0.000	1,531.914	61.310	0.000
f13	581.974	56.145	0.000	608.413	69.615	0.000
f14	1,942.953	218.727	0.000	2,045.749	264.606	0.000
c16	293,793.100	14,460.390	0.000	278,294.200	18,889.550	0.000
c17	696,101.300	23,752.660	0.000	659,890.400	31,147.580	0.000
c19	14,606.140	6,282.412	0.020	13,916.320	8,106.428	0.086
c21	462.806	118.340	0.000	321.669	128.054	0.012
_cons	-4,515,981	4,153,364	0.277	-12,800,000	5,277,214	0.015
R^2			0.531			0.530

(2) 농업순생산 - 5년 표본

	pooled			random effect		
	계 수	표본오차	p-값	계 수	표본오차	p-값
g02	841,080.500	261,451.400	0.001	960,077.800	313,974.900	0.002
g17	0.008	0.001	0.000	0.010	0.001	0.000
g18	0.013	0.003	0.000	0.015	0.004	0.000
g04	7,502.244	134.428	0.000	7,397.673	161.161	0.000
mage	-104,597.800	154,059.000	0.497	171,549.900	197,787.600	0.386
mages	-129.828	1,299.664	0.920	-2,468.412	1,666.931	0.139
f07	1,495.483	72.542	0.000	1,418.419	91.362	0.000
f08	928.993	115.139	0.000	928.945	143.214	0.000
f09	2,066.975	139.204	0.000	1,831.071	177.235	0.000
f12	1,689.916	48.237	0.000	1,548.104	60.292	0.000
f13	464.558	58.464	0.000	460.068	73.110	0.000
f14	1,788.849	213.776	0.000	1,916.263	257.455	0.000
c16	282,867.200	14,246.370	0.000	262,698.200	18,300.090	0.000
c17	590,502.500	25,081.920	0.000	544,772.300	32,455.860	0.000
c19	15,558.240	6,101.714	0.011	14,229.590	7,755.692	0.067
c21	50.935	148.565	0.732	-59.099	158.845	0.710
_cons	2,323,737	4,516,093	0.607	-5,526,718	5,800,747	0.341
R^2			0.557			0.557

(3) 농업상시종사자 1인당 농업소득 - 전체 표본

	pooled			random effect		
	계수	표본오차	z-값	계수	표본오차	z-값
g17p	0.003	0.000	0.000	0.004	0.001	0.000
g18p	-0.010	0.002	0.000	-0.007	0.003	0.016
g04p	4,403.611	142.175	0.000	4,101.201	187.312	0.000
mage	244,375.400	83,487.360	0.003	556,606.100	128,289.600	0.000
mage _s	-2,147.835	700.219	0.002	-4,698.730	1,074.636	0.000
f07p	1,598.233	66.004	0.000	1,483.574	94.453	0.000
f08p	685.845	123.833	0.000	584.726	180.912	0.001
f09p	1,430.542	134.750	0.000	1,080.423	205.816	0.000
f12p	659.524	49.902	0.000	571.917	74.064	0.000
f13p	202.420	60.920	0.001	316.524	87.254	0.000
f14p	1,460.661	229.516	0.000	1,703.539	319.708	0.000
c16p	288,307.700	14,711.380	0.000	233,504.800	22,993.970	0.000
c17p	646,839.700	27,149.780	0.000	577,418.800	42,696.310	0.000
c19p	24,002.520	5,772.255	0.000	15,960.840	8,172.402	0.051
c21p	21.897	91.328	0.811	-110.070	95.447	0.249
_cons	-7,330,784	2,445,859	0.003	-16,400,000	3,765,407	0.000
R ²			0.254			0.252

(3) 농업상시종사자 1인당 농업소득 - 5년 표본

농업소득	pooled			random effect		
	계수	표본오차	z-값	계수	표본오차	z-값
g17p	0.004	0.001	0.000	0.005	0.001	0.000
g18p	-0.022	0.003	0.000	-0.018	0.004	0.000
g04p	4,443.623	151.175	0.000	4,318.738	175.070	0.000
mage	91,013.750	89,360.880	0.308	186,804.900	109,917.600	0.089
mage _s	-1,032.456	751.237	0.169	-1,827.413	923.650	0.048
f07p	1,458.654	76.866	0.000	1,336.267	92.783	0.000
f08p	481.376	130.968	0.000	428.828	156.703	0.006
f09p	1,439.284	137.534	0.000	1,240.235	167.331	0.000
f12p	685.520	50.833	0.000	634.806	61.180	0.000
f13p	161.218	62.605	0.010	187.910	75.222	0.012
f14p	1,361.177	226.669	0.000	1,480.288	266.467	0.000
c16p	275,580.100	14,294.480	0.000	250,941.600	17,446.780	0.000
c17p	548,487.900	29,008.030	0.000	516,289.600	35,848.180	0.000
c19p	23,909.750	5,555.718	0.000	20,487.160	6,570.807	0.002
c21p	-151.493	103.526	0.143	-216.623	104.083	0.037
_cons	-2,005,934	2,612,299	0.443	-4,716,298	3,215,294	0.142
R ²			0.264			0.263

(4) 농업상시종사자 1인당 농업순생산 - 전체 표본

	pooled			random effect		
	계 수	표본오차	z-값	계 수	표본오차	z-값
g17p	0.005	0.000	0.000	0.006	0.001	0.000
g18p	0.009	0.002	0.000	0.006	0.003	0.026
g04p	7,812.127	147.068	0.000	7,194.946	193.752	0.000
mage	21,868.250	86,360.820	0.800	466,613.600	133,456.500	0.000
mages	-633.181	724.319	0.382	-4,326.729	1,117.848	0.000
f07p	1,825.260	68.275	0.000	1,830.039	98.020	0.000
f08p	741.301	128.095	0.000	699.034	187.714	0.000
f09p	2,060.734	139.388	0.000	1,732.413	214.014	0.000
f12p	1,594.622	51.620	0.000	1,325.438	76.889	0.000
f13p	615.873	63.017	0.000	716.696	90.476	0.000
f14p	1,877.283	237.415	0.000	2,152.347	331.218	0.000
c16p	313,549.400	15,217.710	0.000	263,211.800	23,910.610	0.000
c17p	704,135.300	28,084.220	0.000	629,521.300	44,439.190	0.000
c19p	21,270.670	5,970.924	0.000	15,656.670	8,466.054	0.064
c21p	438.948	94.472	0.000	131.904	98.220	0.179
_cons	-585,361	2,530,040	0.817	-13,000,000	3,917,477	0.001
R^2	0.461			0.457		

(4) 농업상시종사자 1인당 농업순생산 - 5년 표본

	pooled			random effect		
	계 수	표본오차	z-값	계 수	표본오차	z-값
g17p	0.007	0.001	0.000	0.008	0.001	0.000
g18p	0.008	0.003	0.021	0.009	0.004	0.019
g04p	7,744.043	155.535	0.000	7,453.750	181.252	0.000
mage	-157,229.900	91,938.290	0.087	-12,346.340	114,226.000	0.914
mages	737.375	772.905	0.340	-501.545	959.818	0.601
f07p	1,548.065	79.083	0.000	1,443.874	96.307	0.000
f08p	500.276	134.746	0.000	475.332	162.551	0.003
f09p	1,909.653	141.501	0.000	1,698.132	173.786	0.000
f12p	1,642.085	52.299	0.000	1,532.798	63.489	0.000
f13p	539.768	64.411	0.000	570.274	78.054	0.000
f14p	1,732.340	233.206	0.000	1,897.331	276.137	0.000
c16p	294,904.600	14,706.780	0.000	272,338.100	18,121.830	0.000
c17p	593,164.200	29,844.710	0.000	560,229.800	37,264.030	0.000
c19p	20,575.330	5,715.960	0.000	17,935.300	6,810.902	0.008
c21p	-68.655	106.512	0.519	-152.222	106.982	0.155
_cons	5,340,883	2,687,645	0.047	1,507,570	3,341,513	0.652
R^2	0.484			0.483		

주 : 변수명은 <부록 2> 168페이지 참조

제 6 장 결 론

1. 시사점

본고는 먼저 일반적인 패널자료의 장단점을 살펴보고 국내 패널조사의 현황과 특징을 정리한 다음 패널조사 분석방법론의 일부를 소개하였다. 또한 통계패키지를 사용하여 패널자료를 분석하는 방법을 예시하였다.

이와 같은 기본 현황을 바탕으로 본 연구에서는 농가경제조사를 패널자료화한 결과와 패널화된 자료를 이용하여 실제로 자료를 분석하는 예를 제시하였다. 본 연구의 결과에 의하면 농가경제조사의 경우 비록 표본을 추적하는 패널조사가 아니라는 점에서 한계가 있지만, 실제로는 패널유지율 면에서 추적 조사하는 일반패널조사에 비해 표본유지율이 낮지 않음을 보여주었다. 이와 같은 결과는 농가경제조사를 패널자료로 만들어서 사용하는 데 큰 제약이 없다는 점을 의미하는 것이라고 하겠다. 또한 통계청의 다른 반복 횡단면 자료의 경우에도 농가패널조사와 같이 적절한 수준의 표본유지율을 달성한 경우라면 패널화를 통해 다양한 이점을 얻을 수 있다는 점을 시사하는 것이라고 하겠다.

반복횡단면 자료를 이용하여 패널데이터를 생성하는 경우 가중치의 문제가 제기된다. 대부분의 경우 패널탈락은 체계적인 요인에 의해 이루어지는 경우가 많으며, 이를 보정하지 않는 경우 자료 분석의 대표성을 확보하기 어렵기 때문이다. 그런데 대부분의 패널자료에서 가중치를 작성하는 경우에 봉착하는 문제가 신규 진입표본의 가중치를 부여하는 방안의 모호성이다. 본고의 분석에 의하면 농가패널조사와 같은 반복 횡단면 자료의 경우 이미 횡단면가중치가 존재하기 때문에 신규 진입표본의 가중치설정 문제가 발생하지 않는다는 점을 알 수 있다. 또한 기존의 횡단면가중치를 이용하여 종단면가중치를 사후 총화할 수 있다는 장점도 있다. 결국 본고의 분석결과는 횡단면가중치가 존재하는 반복횡단면자료의 경우 종단면가중치가 쉽게 도출될 수 있다는 점을 보여주는 것이라고 하겠다.

농가경제조사의 패널화자료를 이용하여 분석한 결과에 의하면 농가의 경우 1분위와 5분위에서는 소득의 고착화현상이 발견되며, 이러한 고착

화현상은 농가소득보다 농업소득에서 더욱 크게 나타났다. 이에 반해 3분위에 속하는 가구의 경우에는 상대적으로 소득이동이 활발하였는데, 상향이동보다는 하향이동의 비중이 높은 것으로 나타났다. 농가자산의 경우에는 농가자산의 분위별 고착화현상이 농업소득이나 농가소득에 비해 월등히 높은 것으로 나타났다. 이외에도 농가부채의 경우 농가부채가 가장 많은 5분위가구의 부채문제가 소폭이나 개선되고 있지만 개선속도는 매우 느린 것으로 나타났다.

또한 한국 농가의 경우 전업농가나 2종 겸업의 경우에는 동일한 형태를 유지하는 비중이 높지만, 1종 겸업의 경우에는 동일한 형태를 유지하는 비중이 상대적으로 낮으며, 전업농가로 변화하는 비중이 높게 나타났다. 전업농가의 경우에도 적지 않은 경우 1종 겸업으로 전환하는 경우가 나타나고 있어서 전업농가와 1종 겸업농가간의 상호이동성이 큰 것으로 나타났다.

임의효과 패널모형을 이용하여 농업소득이나 농업순생산을 결정하는 요인을 분석한 결과에 의하며 다양한 요인들이 통계적으로 유의한 요인으로 판명되었다. 농가자산은 많을수록 농가부채는 적을수록 농업소득이나 농업순생산이 많은 것으로 나타났다. 농업상시 종사자수도 농업소득이나 농업순생산을 높이는 유의한 변수였으며, 차용지이건 자영지이건 간에 논과 밭, 과수원은 모두 유의한 설명변수로 판명되었다. 다만, 다른 조건이 동일한 경우 면적 당 기여도는 과수원이 가장 크게 나타났다. 이외에도 닭을 제외한 가축류도 농업소득이나 농업순생산을 높이는 요인이었으며 마리당 기여도는 젓소가 가장 크게 나타났다.

이와 같은 분석결과를 다시 한번 요약하면 농가경제조사와 같은 반복횡단면 자료의 경우 사후적으로 패널화함으로써 다양한 동태분석이 가능하다. 또한 횡단면가중치의 존재를 감안하였을 때 종단면가중치 문제도 비교적 수월하게 해결할 수 있는 가능성이 있음을 알 수 있다. 본고의 경험을 바탕으로 할 때 통계청에서 실시하고 있는 유사한 구조의 반복 횡단면조사의 경우에도 패널화 작업을 수행할 필요가 있다고 판단된다.

2. 한계와 향후 연구과제

그러나 본고는 다양한 한계점을 가지고 있다. 패널자료분석기법은 다양한 기법 중에 일부를 소개한 것에 불과하며, 패널조사의 현황에 대한 서베이도 전부를 포괄하고 있지는 못하다. 패널자료 분석기법은 관련 교재를 참고할 수 있으며, 패널조사의 현황에 대해서는 이희길 외(2008) 등의 자료를 참고할 수 있다.

한편 본고에서 계산한 종단면가중치는 응답률 모형에 근거하고 있는데, 동 응답률 모형은 모형설정의 오류(mis-specification)의 문제를 가질 수 있다. 향후 보다 다양한 응답률 모형을 이용하여 본고의 결과에 대한 강건성(robustness)을 검토할 필요가 있다. 또한 보다 다양한 기법을 이용하여 가중치를 설정하는 방안도 모색할 필요가 있다. 실증분석에 사용한 농업생산이나 농업소득의 경우에도 모형설정의 문제점을 검토할 필요가 있다.

본고에서 제시한 일부 분석결과는 본격적인 패널데이터의 분석이라기 보다는 분석주제나 분석 방법을 예시한 것이라고 할 수 있다. 본 연구를 통해서 작성된 패널화 데이터를 이용하여 실제적인 정책적 함의점을 찾는 연구나, 정책 성과를 분석하는 연구가 향후에도 지속되어야 할 것이다.

또한 현재의 농가경제조사는 패널화를 염두에 두고 작성되고 있지는 않다. 향후 동 자료가 패널화될 수 있다는 점을 감안하여 설문서를 일부 보강하거나 보다 다양한 분석이 가능하도록 부가조사를 실시하는 방안도 진지하게 검토할 필요가 있다.

참 고 문 헌

- 강석훈(2007), "Introduction to Linear Panel Data Analysis", 한국청소
년개발원 패널분석방법론 세미나 자료
- 이지연(2005), 가구조사 자료의 종단화 방안, 통계청 통계연구보고서
05-03
- 이희길 외(2008), 통계조사의 정확성과 효율성 제고 방안, 연구보고서
2008-04, 통계개발원
- 김용성(2007), "사회·복지 프로그램 패널자료 구축에 관한 타당성 연
구" 경제·인문사회연구회 협동연구 총서 07-03-01
- 백화중 외(2004), "빈곤 및 공공부조 패널 데이터 구축방안" 한국보건사
회연구원 연구보고서(2004-18)
- Clarke P. and P. Tate (1999), "Methodological Issues in the Production
and Analysis of Longitudinal Data from the Labor Force
Survey", Methodology Series No. 17, Office of National Statistics
- Haisken-DeNew(2001). " A Hitchhiker's Guide to the World's
Household Panel Study", DIW Research Note No.1
- Hsiao, C. (2001), Analysis of Panel Data, Cambridge
- Guskova, E., S. Heeringa, K. McGonagle and R. Schoeni (2008),
Panel Study of Income Dynamics : Revised Longitudinal
Weights 1993-2005", Survey Research Center, University of
Michigan

<부록 1> 영국의 노동력조사 패널화 예시

영국에서는 노동력조사(Labor Force Survey)를 패널화하여 노동시장에 대한 동태적 분석을 실시하고 있다. 영국의 노동력조사는 1992년부터 순환패널(rotating panel)형태로 진행되고 있다. 동 조사는 일단 표본에 편입되면 5분기간 연속적으로 조사에 응하게 되며 매 분기마다 1/5의 표본이 교체된다. Clarke and Tate(1999)는 동 자료를 패널화하였을 때 발생하는 문제점에 대하여 연구한 결과를 제시하고 있다. 동 연구에서 10대의 경우에 탈락확률이 높으며 독신의 경우에도 탈락확률이 상대적으로 높은 것으로 나타났다(<부표-1> 참조).

<부표-1> 탈락확률이 높은 무응답자의 특성

번 수	범 주	무 응 답 확 률			
		웨이브1	웨이브2	웨이브3	웨이브4
연령	18-19	1.89	2.56	2.86	1.92
	20-24	1.79	2.08	2.1	2.83
	25-29	1.17	1.3	1.44	1.55
거주형태	민간임대	2.12	1.52	1.86	2.29
결혼여부	독신	1.25	1.12*	1.27	1.49
경제활동상태	전시간,임시근로자	1.12*	1.36*	1.13*	1.75

자료 : Clarke and Tate(1999), Table1

주 : *는 5%에서 유의하지 않은 경우

〈부표-2〉 초기 가중치를 이용한 분석결과

변수		횡단면수 (단순)	횡단면 (가중)	패널화 (단순)	패널화(가중)		
					A	B	C
거주형태	자가	73.7	73	76	73	72.9	73
	공공임대	17.3	17.5	16.7	17.6	17.6	17.6
	민간임대	9	9.5	7.3	9.5	9.5	9.5
결혼상태	독신	24.9	26.3	22.8	26.5	25.4	26.5
	결혼	58.2	56.8	60.8	56.5	57.8	56.5
	동거	8.3	8.6	8.1	8.8	8.7	8.8
	별거/이혼 등	8.5	8.4	8.4	8.2	8.1	8.2
경제활동 상태 (분기1)	고용	72.2	72.3	73.2	73	72.4	72.5
	ILO실업	6.9	7.2	6.4	7	7.3	7.2
	비경제활동	20.9	20.6	20.4	20	20.3	20.3
경제활동 상태 (분기2)	고용	71.7	72.2	73.6	73.4	73	73
	ILO실업	6.6	6.8	5.9	6.4	6.6	6.6
	비경제활동	21.2	21	20.5	20.2	20.4	20.4

자료 : Clarke and Tate(1999), Table2

주 : A는 거주형태, B는 거주형태+경제활동상태, C는 거주형태+결혼상태+경제활동상태

실제로 영국의 5분기 노동력조사의 패널화 데이터는 다음과 웹사이트에서 일반인들에게 공개되어 있다.

<http://www.esds.ac.uk/findingData/snDescription.asp?sn=5660>

<부록 2> 농업소득 및 농업순생산 결정요인분석에서 사용한 변수의 기초통계량

A. 가구 농업소득 및 농업순생산

① 전체 표본

전체표본	설 명	관측치수	평 균	표준편차
g02	농업상시종사자수	15,985	1.8	0.7
g08	농업소득(천원)	15,985	12,090.00	22,237.70
g19	농업순생산(천원)	15,985	19,122.80	27,772.70
g17	농가자산(천원)	15,985	301,210.80	356,506.90
g18	농가부채(천원)	15,985	31,131.50	63,737.60
g04	차영업노동시간	15,985	1,522.70	1,579.50
mage	가구주연령	15,985	60.7	10.6
msex	가구주 성별	15,974	1.1	0.2
f07	자작지 논면적	15,974	1,498.80	2,435.40
f08	자작지 밭면적	15,974	788.6	1,504.40
f09	자작지 과수원 면적	15,974	381.3	1,175.90
f12	차용지 논면적	15,974	1,333.50	3,364.60
f13	차용지 밭면적	15,974	978.9	2,983.80
f14	차용지 과수원 면적	15,974	104.1	698.5
c16	한우	15,974	2.2	10.7
c17	젓소	15,974	0.7	6.5
c19	돼지	15,974	4.5	24.1
c21	닭	15,974	55	1,337.40

② 5년 유지표본

5년표본	설 명	관측치수	평 균	표준편차
g02	농업상시종사자수	12,710	1.8	0.7
g08	농업소득(천원)	12,710	12,756.00	21,046.70
g19	농업순생산(천원)	12,710	20,069.70	26,701.60
g17	농가자산(천원)	12,710	293,269.40	327,654.30
g18	농가부채(천원)	12,710	30,946.70	56,610.70
g04	자영업노동시간	12,710	1,585.70	1,536.80
mage	가구주연령	12,710	60.6	10.4
msex	가구주 성별	12,710	1	0.2
f07	자작지 논면적	12,710	1,570.60	2,383.90
f08	자작지 밭면적	12,710	810	1,472.90
f09	자작지 과수원 면적	12,710	411.4	1,233.50
f12	차용지 논면적	12,710	1,449.50	3,577.10
f13	차용지 밭면적	12,710	1,031.70	3,042.60
f14	차용지 과수원 면적	12,710	115	751.3
c16	한우	12,710	2.5	11.5
c17	젓소	12,710	0.7	6.5
c19	돼지	12,710	4.8	26
c21	닭	12,710	35.7	1,068.50

170 농가소득의 동태적 변화요인

B. 농업상시종사자 1인당 농업소득 및 농업순생산

① 전체 표본

전체표본	설 명	관측치수	평 균	표준편차
per_inc	1인당 농업소득(천원)	15,251	6,620.20	12,559.00
per_pro	1인당 농업순생산(천원)	15,251	10,514.60	15,269.20
g17p	1인당 농가자산(천원)	15,251	175,702.50	235,877.90
g18p	1인당 농가부채(천원)	15,251	17,691.30	41,632.20
g04p	1인당 자영업노동시간	15,251	837.1	757.9
f07p	1인당 자작지 논면적	15,240	855.8	1,438.00
f08p	1인당 자작지 밭면적	15,240	435.2	746.3
f09p	1인당 자작지 과수원 면적	15,240	217.1	708.4
f12p	1인당 차용지 논면적	15,240	750.9	1,890.50
f13p	1인당 차용지 밭면적	15,240	539.4	1,589.70
f14p	1인당 차용지 과수원 면적	15,240	58.2	388.4
c16p	1인당 한우	15,240	1.2	6.1
c17p	1인당 젓소	15,240	0.4	3.3
c19p	1인당 돼지	15,240	2.7	15.3
c21p	1인당 닭	15,240	34.7	989.3

② 5년 표본

5년표본	설 명	관측치수	평 균	표준편차
per_inc	1인당 농업소득(천원)	12,215	6,885.30	11,719.30
per_pro	1인당 농업순생산(천원)	12,215	10,873.10	14,400.50
g17p	1인당 농가자산(천원)	12,215	167,366.70	208,462.90
g18p	1인당 농가부채(천원)	12,215	17,291.60	31,686.10
g04p	1인당 자영업노동시간	12,215	859.3	750.4
f07p	1인당 자작지 논면적	12,215	876.3	1,285.50
f08p	1인당 자작지 밭면적	12,215	441.3	731.5
f09p	1인당 자작지 과수원 면적	12,215	231.5	738.4
f12p	1인당 차용지 논면적	12,215	798.3	1,939.60
f13p	1인당 차용지 밭면적	12,215	559.8	1,618.80
f14p	1인당 차용지 과수원 면적	12,215	63.2	408.8
c16p	1인당 한우	12,215	1.3	6.6
c17p	1인당 젓소	12,215	0.4	3.2
c19p	1인당 돼지	12,215	2.8	16.5
c21p	1인당 닭	12,215	23.1	883.9

<부록 3> 외국의 패널조사 현황(자료: PSID 웹 사이트)

A. 오스트레일리아

- ① 제목: Household Income and Labor Dynamics in Australia(HILDA)
- ② 주관기관: University of Melbourne,
Melbourne Institute of Applied Economic and Social Research
- ③ 조사연도: 2001
- ④ 표본: 7,682가구의 구성원들에게 인터뷰 실시.
- ⑤ 내용: 소득사항, 노동시장 관련사항, 가족사항
- ⑥ Site 주소: <http://www.melbourneinstitute.com/hilda>

B. 벨기에

- ① 제목: Belgian Socio-economic panel(SEP)
- ② 주관기관: Centre for Social Policy, University of Antwerp
- ③ 조사연도: 1985, 1988, 1992, 1997
- ④ 표본: 표본은 벨기에 가구와 개인을 대표함. 1985년 6,471, 1997년 4,632(신규 표본 2,375가구 포함)
- ⑤ 내용: 인구사회학적 특성, 고용상황, 가구소득, 자산, 주거상황, 주관적 소득 평가
- ⑥ Site 주소: http://www.ufsia.ac.be/CSB/sep_nl.htm

C. 캐나다

- ① 제목: Survey of Labor Income Dynamics(SLID)
- ② 주관기관: Statistics Canada
- ③ 조사연도: 1993~2000
- ④ 표본: 10개주(province)의 약 35,000가구
- ⑤ 내용: 경제복지상황, 복지의 영향과 변화, 교육, 장애
- ⑥ Site 주소: <http://www.statcan.ca/english/survey/household/dynamic/income.htm>

D. 네덜란드

- ① 제목: Dutch Socio-Economic Panel(ISEP)
- ② 주관기관: Statistics Nrtherlands
- ③ 조사연도: 1984~1997
- ④ 표본: 5,000가구의 16세 이상의 가구구성원
- ⑤ 내용: 교육, 소득, 자산, 적응, 복지에 대한 평가 등
- ⑥ Site 주소: <http://center.uvt.nl/research/facilities/esp.htm>

E. 프랑스

- ① 제목: French Household Panel
- ② 주관기관: ADEPS, Equipe de recherche en Analyse Dynamique des Effets dea Politiques Sociales and Direction Regionale en Lorraine de l'Institut National de la Sytistique et des Etudes Economiques(INSEE)
- ③ 조사연도: 1985~1990
- ④ 표본: 최초 표본은 Lorraine 지방에 사는 715가구였으며, 두 번째 웨이브에서 2,092가구로 증가
- ⑤ 내용: 가구구성, 인구사회학적 특성, 주거, 소득, 교육, 고용, 주거 배경, 주관적 빈곤지수, 고용창출 활동, 가구자산, 부채 등
- ⑥ Site 주소: <http://www.ceps.lu/paco/pacofrpa.htm>

F. 룩셈부르크

- ① 제목: Panel Socio-Economique "LiewenZu LetZëbuerg"(PSELL)
- ② 주관기관: A Centre for Socio-Economic Research
- ③ 조사연도: 1985-
- ④ 표본: 1985~94/ 2,012가구와 6,110명의 개인, 1994~/ 2,978가구와 8,232명의 개인
- ⑤ 내용: 주거, 가족구성, 소비수준, 부채관리, 빈곤측정 등
- ⑥ Site 주소: <http://www.ceps.lu/psell/pselpres.htm>

G. 러시아

- ① 제목: Russia Longitudinal Monitoring Survey(RLMS)
- ② 주관기관: Carolina Population Center at the University of North Carolina
- ③ 조사연도: 1992-
- ④ 표본: 8,200가구
- ⑤ 내용: 건강, 식품섭취, 지출의 측정, 서비스 이용, 지역사회 수준 등
- ⑥ Site 주소: <http://www.cpc.unc.edu/projects/rlms/home.html>

H. 스웨덴

- ① 제목: The Swedish Panel Study Market and Nonmarket Activities(HUS)
- ② 주관기관: Department of Economics, Gothenburg University
- ③ 조사연도: 1984
- ④ 표본: 시설에 거주하지 않는 19~74세의 2,000명
- ⑤ 내용: 가족형태, 아동보호, 주거, 고용, 소득, 재산, 세금, 서비스 이용 등
- ⑥ Site 주소: <http://www.nek.uu.se/faculty/klevmark/hus.htm>

I. 스위스

- ① 제목: Swiss Household Panel(SHP)
- ② 주관기관: SHP-Team at the University of Neuchâtel
- ③ 조사연도: 1999-
- ④ 표본: 5,074가구의 7,79명의 개인
- ⑤ 내용: 가구구성, 주거, 소득, 가족생활, 주요사건, 가족 이슈, 건강, 교육, 고용, 종교, 네트워크, 레저, 미디어, 정치적 가치, 만족도
- ⑥ Site 주소: <http://www.swisspanel.ch/>

J. 영국-ECHP

- ① 제목: European Community Household Panel Study(ECHP)
- ② 주관기관: The London School of Economics and Political Science(LSE)
- ③ 조사연도: 1994~1996
- ④ 표본: European Union Member States에서 가구패널 표본구성
- ⑤ 내용: 이민, 주거, 소득, 경제활동, 소비, 사회적 관계, 교육, 건강 등
- ⑥ Site 주소: <http://www.iser.essex.sc.uk/bhps/index.php>

K. 미국-HRS

- ① 제목: Health and Retirement Study(HRS)
- ② 주관기관: The Institute For Social Research at the University of Michigan
- ③ 조사연도: 1992~
- ④ 표본: 원표본은 1931~41년에 태어난 대상, 이후 1923년 이전에 태어난 사람, 1998년 웨이브에서 1924~30년, 1942~47년에 태어난 사람 더해짐, 2004년에 1948~53년에 태어난 사람이 더해짐.
- ⑤ 내용: 소득과 고용, 소득과 부, 건강상태, 건강보험의 커버정도, 세대간 이전, 가족구성 등
- ⑥ Site 주소: <http://www.umich.edu/~hrswww/>

◆ 집필자

- 방태경(통계개발원 사회통계실 사무관)
- 강석훈(성신여자대학교 경제학과 교수)

농가소득의 동태적 변화 요인
-농가경제조사 자료의 패널화와 그분석-

· 발행연월일	2008년 12월 15일 인쇄 2008년 12월 26일 발행
· 발 행 인	이재형
· 발 행 처	통계개발원 ☎ 02-8228-3002 대전광역시 서구 한밭대로 614 (둔산동 979) 신협건물 13층 ☎ 대표 (042) 717-0203 Fax (042) 717-0251