# 한국노동패널 사례보고

장 지 연*

**<요 약>**

한국노동패널조사(Korea Labor and Income Panel Study)는 비농촌지역에 거주하는 한국의 가구와 가구원을 대표하는 패널표본구성원(5,000가구에 거주하는 가구원)을 대상으로 1년에 1회 경제활동 및 노동시장 이동, 소득 활동 및 소비, 교육 및 직업훈련, 사회생활 등에 관하여 추적 조사하는 종단면 조사(longitudinal survey)이다. 매년 반복적인 조사를 통해 표본가구구성원들에 대한 다양한 내용의 정보를 수집하는 가운데 조사년수가 늘어가면서 표본가구원 개인들의 생애과정(Life Course)에 걸친 단계별 변화와 이동의 이력(履歷), 특히 학교교육력(Schooling History), 취업력(Work History) 혹은 노동시장 이동(Labor Market Transitions) 과정이 구축될 수 있을 것이다.

조사구의 추출방법은 계통추출방법을 사용하였으며, 제주도를 제외한 전국의 시부만을 대상으로 1,000개의 조사구를 선정하고, 각 조사구내에서는 '97년 고특의 조사대상가구 중에서 5가구를 임의 선정(random sampling)하였다.

## Ⅰ. 한국노동패널의 목적

한국노동패널조사(Korea Labor and Income Panel Study)는 비농촌지역에 거주하는 한국의 가구와 가구원을 대표하는 패널표본구성원(5,000가구에 거주하는 가구원)을 대상으로 1년에 1회 경제활동 및 노동시장 이동, 소득 활동 및 소비, 교육 및 직업훈련, 사회생활 등에 관하여 추적 조사하는 종단면 조사(longitudinal survey)이다.

*연락처: 서울시 영등포구 여의도동 16-2 중소기업회관빌딩 9층 한국노동연구원
전화번호: 784-9339, fax: 784-7684, e-mail : jchang@kli.re.kr
자료정리: 한국노동연구원 노동패널 연구팀

현재 우리 나라에서 실시되고 있는 노동시장관련 전국규모 조사로는 통계청의 「경제활동인구조사」와 「고용구조특별조사」, 노동부의 「노동력유동실태조사」 및 「임금구조기본통계조사」 등이 있다. 그러나 이상의 통계조사들은 모두 일정시점을 기준으로 한 횡단면적(cross-sectional) 조사로서 노동시장의 정지된 그림(snapshot view)을 보여줄 수는 있으나 동적인 측면에서 개인의 경제활동참여 및 노동시장 이동과정, 소득과 소비의 변화 등을 파악하는 데는 한계가 있었다.

패널조사는 횡단면 조사의 이러한 단점을 보완하면서 동적인 차원에서 가구와 개인의 장기간에 걸친 변화와 상태간 이동과정을 보여줌으로써 횡단면 자료만 가지고는 불가능한 심도 있는 정책연구 및 정책평가를 가능하게 해주는 장점을 가지고 있다. 학교를 졸업한 청소년들의 노동시장 진입 및 이동과정, 실업자들의 재취업 과정 및 장기실업의 동태, 실업급여의 재취업효과, 개인의 의식구조와 노사관계의 변화, 육아장려금과 주부의 경제활동변화 등에 관한 연구들은 패널자료의 분석을 통하여서만이 가능한 많은 중요한 노동시장관련 정책연구 과제들 중의 일부에 지나지 않을 뿐이다.

선진국들에서는 이러한 정책연구의 필요성을 충족시키고 노동시장 기초연구의 활성화를 위해 우리보다 앞서 패널조사를 시작하여 오랜 기간 성공적으로 수행해 오고 있다. 미국의 예를 들면 노동시장에 초점을 맞추어 1966년부터 시작된 NLS(National Longitudinal Survey)와 1979년부터 시작된 NLSY(National Longitudinal Survey of Youth), 그리고 가구와 개인의 소득 및 소비에 초점을 맞추어 1968년에 시작된 PSID(Panel Study of Income Dynamics) 등은 20년 이상 조사가 장기 지속됐거나 현재도 진행중인 대표적인 패널조사로 꼽힌다. 그 외에도 1993년에 시작된 캐나다의 SLID(Survey of Labor and Income Dynamics), 1991년에 시작된 BHPS(British Household Panel Study), 1984년에 시작된 독일의 GSEP(German Socio-Economic Panel) 등이 있다.

OECD, ILO 등 국제기구에 이미 가입하고 있고 선진국으로의 진입을 눈앞에 둔 현재 우리 나라는 급격한 사회·경제적 변화와 함께 노동시장도 많은 질적·양적 변화를 겪고 있으며, 이에 따라 보다 구체적이고 미시적인 노동정책에 대한 수요가 증대하고 있다. 올바른 정책의 수립 및 실행, 평가를 위해서는 대표성 있고 신뢰할만한 기초자료의 뒷받침이 필수적이지만 현재 우리 나라에서 실시하고 있는 각종 노동시장관련 통계조사들은 아직 그 내용이 빈약하거나 개별수준이 아닌 기업수준의 조사들이어서 개인의 특성 및 의사결정에 바탕을 둔 미시적이고 동적인 차원에서의 노동시장의 이해 및 정책결정에 별다른 도움을 주지 못하고 있는 실정이다. 동시에 경제와 노동시장의 국제화가 급속히 진행되고 있는 가운데 다른 나라들과의 비교분석이 가능한 노동시장자료 또한 부재한 상황이다.

이러한 시점에 한국노동패널조사는 노동시장 관련 기초조사자료의 미흡 및 부재의 문제를 해소하고 나아가 분석적인 노동시장연구를 활성화시킴으로써, 보다 합리적이고 정확한 노동시장 및 고용정책의 수립과 평가에 기여하고자 1998년 처음 시작되었다. 앞으로 노동패널데

이터가 장기간에 걸쳐 성공적으로 구축될 경우 특별히 각종 노동정책의 수립 및 평가, 고용보험제도의 효율적인 운영과 장기적인 발전에 기여할 뿐만 아니라 노동관련 연구의 활성화를 위한 귀중한 기초자료로 자리잡을 것으로 기대된다.

한국노동패널조사와 연구는 노동부의 고용정책에 의해 배정되는 고용보험 연구사업비에서 예산 지원을 받고 있으며, 아울러 본 조사는 통계청의 공식 승인을 받아 수행되고 있음을 밝힌다(통계청 승인번호 : 33301).

## Ⅱ. 조사의 개요(Survey Design)

한국노동패널조사는 조사대상 표본가구와 가구 구성원들에 대하여 년 1회 반복적인 면접조사를 통하여 관련 정보를 수집한다. 조사의 내용은 가구에 관련된 내용과 개인에 관련된 내용으로 나뉘어지는데 가구와 가구구성에 관한 사항은 가구를 대표하는 가구주 혹은 가구주의 배우자가 주응답자가 되며, 개인용 조사에서는 조사당일 기준 만 15세 이상의 가구원들을 대상으로 그들의 경제활동, 소득 및 소비, 노동시장이동, 학교교육 및 직업훈련 등과 관련된 다양한 정보를 수집한다.

본 조사의 1차년도(1998년) 조사대상에 포함된 원(原)표본가구는 제주도를 제외한 전국 비농촌지역에 거주하는 가구를 대표하도록 표집이 되었으며 따라서 원표본가구원들(패널멤버들)은 우리나라 비농촌지역의 거주인구를 대표한다. 2차년도 조사부터는 1차년도 원표본 가구원들에 대한 연례조사가 수행되며 이사 및 분가한 경우는 추적조사가 이루어진다. 원표본가구원으로부터 출생한 자녀들은 표본가구원으로 추가되어 조사 및 추적의 대상이 되며, 원표본가구원, 혹은 그의 자녀가 결혼 등을 통하여 배우자를 맞이한 경우 그 배우자도 혼인관계가 지속되는 한 조사의 대상에 포함되며, 그 반대로 원표본가구로부터 분가하여 단독으로 혹은 배우자와 함께 새로운 가구를 형성할 경우 그 가구는 분가표본가구가 되며 원표본가구원 뿐만 아니라 그 배우자도 조사의 대상에 포함된다. 이러한 추적의 원칙에 따라서 조사년도가 증가하면서 조사대상 가구 및 가구원은 가구원 모두의 사망이나 이민 등으로 표본에서 탈락하는 경우를 제외하고는 확대생성되는 과정을 거치게 된다. 이러한 조사설계의 목적은 원표본가구 및 가구원들을 통해 이루어지는 가구형성과 해체의 다이나믹스가 조사대상표본에 반영되면서 동시에 어느 시점에서나 원표본가구와 가구원의 대표성이 유지되도록 하는데 있다.

한국노동패널에서는 매년 반복적인 조사를 통해 표본가구구성원들에 대한 다양한 내용의 정보를 수집하는 가운데 조사년수가 늘어가면서 표본가구원 개인들의 생애과정(Life Course)에 걸친 단계별 변화와 이동의 이력(履歷), 특히 학교교육력(Schooling History), 취업력(Work History) 혹은 노동시장이동(Labor Market Transitions) 과정이 구축될 수 있을 것이다.

# Ⅲ. 표본 및 표집 방법

## 1. 표본틀 및 추출방법

한국노동패널조사는 '95년 인구주택총조사의 10% 표본조사구(전국 21,938조사구)를 모집단으로 하고 있다. 다만 '95년 인구주택총조사의 10% 표본조사구 중에서 5,000가구를 직접 추출하는 작업이 상당한 시간과 비용을 발생시키기 때문에 시간과 비용을 절약하고, 동시에 '97년 고용구조특별조사의 결과와의 상호비교를 위하여 추출된 표본이 '97년 고용구조특별조사의 표본에 속하도록 하였으며 다음과 같은 단계를 거쳤다.

우선 '97년 고용구조특별조사는 전국을 38개 독립추계집단으로 층화하였는데 비해 한국노동패널조사는 표본수의 한계로 인하여 지역별로 구별하기보다는 제주도를 제외한 전국의 도시지역을 하나의 추계집단으로 하였다. '95 인구주택총조사의 10% 표본조사구(21,938개) 중 제주도를 제외한 전국의 도시지역만을 대상으로 이를 층화하였다. 층화의 기준은 지역을 우선으로 하고, 동일 지역 내에서는 '97년 고특(고용구조특별조사)의 층화방법을 채택하였다.

조사구의 추출방법은 계통추출방법을 사용하였으며, 제주도를 제외한 전국의 시부만을 대상으로 1,000개의 조사구를 선정하고, 각 조사구내에서는 '97년 고특의 조사대상가구 중에서 5가구를 임의 선정(random sampling)하도록 원칙을 정하였다. 조사구의 계통추출시 추출된 조사구가 '97 고용구조특별조사의 조사구에 포함되지 않는 경우에는 가장 가까운 고용구조특별조사의 조사구에 포함되도록 조정하였다. 조사구의 추출과정에서 서울지역과 경기도 시부의 읍면지역은 추출되어야 할 조사구의 숫자보다 '97년 고특의 조사구수가 적기 때문에 해당되는 모든 조사구를 포함시키고 조사구당 가구수를 5가구에서 6가구로 상향조정하는 방식을 택하였다. 이에 따라, 전체의 조사구수는 계획한 1,000개에서 951개로 감소하였으나 지역별 표본가구수는 변동이 없도록 하였다. '95년 인구주택조사의 10% 표본을 기준으로 특정 조사구가 추출될 확률은 서울과 경기도 시부의 읍면지역을 제외하고는 전국적으로 지역과 상관없이 5.3%임. 서울의 경우에는 추출확률 4.4%, 경기도 시부의 읍면 지역은 추출확률 3.5%이다.

위와 같은 방법으로 조사구를 선정한 후 각 조사구내에서 5~6가구를 선정하였는데 선정방법은 '97년 고용구조특별조사에서 조사 성공된 가구들 중 임의추출방법을 이용하여 5~6가구를 선정하였다. 각 조사구내에서 특정 가구가 추출될 확률은 조사구내의 총가구수, '97년 고특에서 성공한 가구수, KLIPS에서 추출한 가구수에 따라 결정된다.

<표 1>

| 시 도 명 | 모집단 조사구수 | 고특 조사구수 | 노동패널 조사구수 |
|---|---|---|---|
| 서 울 | 5,186 | 227 | 227 |
| 부 산 | 1,841 | 176 | 97 |
| 대 구 | 1,212 | 155 | 64 |
| 인 천 | 1,116 | 150 | 59 |
| 광 주 | 607 | 106 | 32 |
| 대 전 | 592 | 101 | 31 |
| 울 산 | 456 | 70 | 24 |
| 경 기 | 3,624 | 280 | 167 |
| 강 원 | 725 | 165 | 26 |
| 충 북 | 694 | 159 | 22 |
| 충 남 | 865 | 122 | 25 |
| 전 북 | 967 | 170 | 40 |
| 전 남 | 996 | 152 | 25 |
| 경 북 | 1,372 | 129 | 54 |
| 경 남 | 1,878 | 290 | 58 |
| 전 체 | 21,938 | 2,497 | 951 |

## 2. 표본 선정과 관련된 기준

조사대상가구 중 실사 진행 과정에서 다음에 해당되는 경우는 해당 조사구 내에서 무작위로 표집된 다른 가구로 대체되었다.

1. 표본의 대체

1) 주소불명인 경우(주소가 기록상으로는 있지만 실제로는 없는 경우)는 대체한다.

2) 조사대상자가 거주지에서 지난 1년 동안 한 번도 살았던 적은 없고 주소만 올려놓은 경우는 대체한다

3) 일반 가구가 아닌 경우(종교시설, 기숙사, 공장 등)는 대체한다(예를 들어 조사대상에 절(寺)이 표집되어 순수한 종교활동을 하는 승려만이 거주하는 경우).

4) 가구주와 배우자 중 한명이라도 응답하지 않은 경우는 대체한다.

5) 가구원 중 설문에 응해야 할 경제활동인구(취업자)가 한 명이라도 빠진 경우는 대체한다.

2. 이사 추적

조사대상자가 거주지에서 지난 1년 동안 살았던 적이 있었고 현재는 다른 곳으로 이사를 간 경우 조사대상자를 추적한다.

# Ⅳ. 1차년도 조사 진행 과정과 조사 결과

## 1. 조사방법

조사방법은 면접타개식 조사방법으로, 면접원이 조사대상자의 가구를 방문하여 직접 조사표를 작성하는 방법을 선택하였다. 특히 가구용 설문지의 경우는 가구에 유치하거나 비가구원이 대리하는 일이 없이 가구원이 반드시 응답을 하도록 하였다. 그러나 개인용 설문지의 경우 해당 개인에 대한 면접 타계식 조사를 원칙으로 하되, 응답자를 만날 수 없거나 응답을 거부할 때는 설문지를 가구에 유치하여 응답자가 직접 조사표에 기입하는 방법을 보충적으로 사용하였다.

## 2. 조사 진행 과정

1차년도 한국노동패널의 실제 조사 기간은 약 5개월 간으로, 1998년 6월 2일에 시작되어 1998년 10월 13일에 완료되었다. 데이타 구성을 위한 부호화(coding) 과정과 데이타 클리닝 과정 등 자료처리 기간은 1998년 6월 8일부터 1998년 11월 27일에 걸쳐 조사와 같이 진행되었다.

조사 진행 과정 동안 자료의 질을 높이기 위해 2단계의 자료 정리 과정을 거쳤다. 첫째, 조사 과정 동안 함께 면접원이 수거해 온 모든 설문지에 대해 교육받은 연구원들이 검증을 실시하였다. 그 과정을 통해 무응답 항목이나 설문 내용을 잘못 이해하고 답한 경우, 설문지 내용 간에 일치성이 없는 경우는 모두 다시 응답자에게 질문토록 하였다. 둘째, 입력해 온 자료에 대해 설문지의 내용과 입력된 내용이 동일한지 검토하였다. 그 결과 자료 입력 과정에서 잘못 입력된 내용은 올바른 내용으로 수정되어 입력하였다.

| | | |
|---|---|---|
| 총조사표본수 | → | 5000가구, 총인원은 17,505명<br>15세 이상 가구원 13,738명 중 면접에 성공한 가구원은 13,317명 |
| 총투입 면접원수 | → | 129명 |
| 실사 기간 | → | 1998년 6월 2일 ~ 1998년 10월 13일 |
| 원표본 성공률 | → | 원표본 5,000가구 중 75.3% 성공률 : 3,773가구<br>(추가 표본은 24.7%로 1,227가구) |
| 자료처리 기간 | → | 1998년 6월 8일 ~ 1998년 11월 27일 |

### 3. 원표본 성공률과 표본 대체율

1. 실사 결과

① 원표본 성공률

원표본 5,000가구 중 75.5%(3,773가구)의 성공률을 보였다. 나머지 24.5%(1,227가구)는 대체하여 성공하였다.

② 5,000가구의 구성

㉮ 5,000가구의 총가구원수는 17,505명

| 가구당 가구원수 | 가구수 | 비 율(%) |
|---|---|---|
| 1 | 517 | 10.3 |
| 2 | 655 | 13.1 |
| 3 | 965 | 19.3 |
| 4 | 1,892 | 37.8 |
| 5 | 696 | 13.9 |
| 6 | 205 | 4.1 |
| 7 | 58 | 1.2 |
| 8 | 9 | 0.2 |
| 9 | 3 | 0.1 |
| 전 체 | 5,000 | 100.0 |

㉯ 15세 이상 가구원은 13,738명. 그 중 응답한 사람은 13,317명

㉰ 13,738명의 취업상태

(가) 임금근로자 : 4,012명

(나) 비임금근로자 : 2,415명

(다) 미취업자 : 7,311명(무응답 : 421명)

㉮ 5,000가구의 지역별 분포

| 지 역 | 조사가구수 | 지 역 | 조사가구수 |
|---|---|---|---|
| 서 울 | 1,362 | 경 기 | 853 |
| 부 산 | 485 | 강 원 | 130 |
| 대 구 | 320 | 충 청 | 235 |
| 인 천 | 295 | 전 라 | 325 |
| 광 주 | 160 | 경 북 | 270 |
| 대 전 | 155 | 경 남 | 290 |
| 울 산 | 120 | 전 국 | 5,000 |

2. 실사진행 상황

(1) 응답 시기

응답 시기는 가구용과 각각의 개인용에서 차이가 있다. 그 이유는 우선 가구용 설문지를 면접원이 가구원과 만나서 가구용 설문지를 작성 한 이후에 개인용 설문지의 조사 대상 가구원을 면접하거나 면접하기 어려운 경우에는 유치를 해서 본인이 작성하도록 했기 때문이다. 이때, 개인이 설문지 작성을 미루거나 거부할 경우, 응답 시기가 가구용 설문지와 달라지게 되고, 한 가구 내에서도 개인간의 응답 시기 차이가 날 수 있다.

① 가구용 설문지 응답시기

| 월 | 가구수 | 비 율(%) |
|---|---|---|
| 6월 | 1,245 | 24.9 |
| 7월 | 1,366 | 27.3 |
| 8월 | 1,246 | 24.9 |
| 9월 | 593 | 11.9 |
| 10월 | 31 | 0.6 |
| 무응답 | 519 | 10.4 |
| 전 체 | 5,000 | 100.0 |

② 개인용 설문지 응답시기

(단위 : 명, %)

| 월 | 개인용 설문지 | 임금근로자 | 비임금근로자 | 미취업자 |
|---|---|---|---|---|
| 6월 | 3,391( 24.7) | 1,004( 25.0) | 599( 24.8) | 1,788( 26.0) |
| 7월 | 3,937( 28.7) | 1,157( 28.8) | 760( 31.5) | 2,020( 29.3) |
| 8월 | 3,576( 26.0) | 1,073( 26.7) | 640( 26.5) | 1,863( 27.0) |
| 9월 | 1,686( 12.3) | 547( 13.6) | 278( 11.5) | 861( 12.5) |
| 10월 | 98( 0.7) | 35( 0.9) | 15( 0.6) | 48( 0.7) |
| 무응답 | 1,050( 7.6) | 196( 4.9) | 123( 5.1) | 310( 4.5) |
| 전 체 | 13,738(100.0) | 4,012(100.0) | 2,415(100.0) | 6,890(100.0) |

(2) 응답자 분포

① 가구용 설문지 응답자

| 가구주와의 관계 | 가 구 | 비 율(%) |
|---|---|---|
| 조부모 | 2 | - |
| 부모 | 59 | 1.2 |
| 가구주 | 2,282 | 45.6 |
| 배우자 | 2,342 | 46.8 |
| 자녀/자녀의 배우자 | 275 | 5.5 |
| 형제자매/형제자매의 배우자 | 28 | 0.6 |
| 손자녀/손자녀의 배우자 | 5 | 0.1 |
| 기타 | 7 | 0.1 |
| 전 체 | 5,000 | 100.0 |

② 개인용 설문지 응답형식

| | 본 인 | 대 리 | 전 체 |
|---|---|---|---|
| 면접 | 7,382 | 1,970 | 9,352 |
| 유치 | 2,585 | 404 | 2,989 |
| 전화 | 143 | 62 | 205 |
| 면접+전화 | 239 | 99 | 338 |
| 유치+전화 | 351 | 82 | 433 |
| 전 체 | 10,700 | 2,617 | 13,317 |

③ 임금근로자용 설문지 응답형식

| | 본 인 | 대 리 | 전 체 |
|---|---|---|---|
| 면접 | 1,827 | 603 | 2,430 |
| 유치 | 1,060 | 123 | 1,183 |
| 전화 | 60 | 23 | 83 |
| 면접+전화 | 99 | 38 | 137 |
| 유치+전화 | 150 | 29 | 179 |
| 전 체 | 3,196 | 816 | 4,012 |

④ 비임금근로자용 설문지 응답형식

| | 본 인 | 대 리 | 전 체 |
|---|---|---|---|
| 면접 | 1,506 | 292 | 1,798 |
| 유치 | 387 | 67 | 454 |
| 전화 | 15 | 10 | 25 |
| 면접+전화 | 49 | 15 | 64 |
| 유치+전화 | 56 | 18 | 74 |
| 전 체 | 2,013 | 402 | 2,415 |

⑤ 미취업자용 설문지 응답형식

| | 본 인 | 대 리 | 전 체 |
|---|---|---|---|
| 면접 | 4,049 | 1,075 | 5,124 |
| 유치 | 1,138 | 214 | 1,352 |
| 전화 | 68 | 29 | 97 |
| 면접+전화 | 91 | 46 | 137 |
| 유치+전화 | 145 | 35 | 180 |
| 전 체 | 5,491 | 1,399 | 6,890 |

3. 대체가구

(1) 지역별 대체율

대체율은 평균 24.5%로 1,227가구를 대체하였다. 지역별로는 서울이 30.2%로 가장 대체율이 높았다. 그리고 가장 대체율이 낮은 지역은 충청으로 16.2%에 달하는 것으로 나타났다.

| 지 역 | 대체율(%) | 지 역 | 대체율(%) |
|---|---|---|---|
| 서 울 | 30.2 | 경 기 | 27.7 |
| 부 산 | 19.4 | 강 원 | 27.7 |
| 대 구 | 25.3 | 충 청 | 16.2 |
| 인 천 | 23.4 | 전 라 | 22.5 |
| 광 주 | 16.3 | 경 북 | 24.1 |
| 대 전 | 18.1 | 경 남 | 16.9 |
| 울 산 | 16.7 | 전 국 | 24.5 |

(2) 대체사유별 대체율

대체를 하게 된 이유를 살펴보면 전체 24.5%의 대체 중 가장 높은 비중을 차지하는 것이 강력 거절로 11.8%, 591가구에 달하는 것으로 나타났다.

| 대체사유 | 대체율(%) |
|---|---|
| 주소불명 | 1.7 |
| 이사→추적 불가 | 6.1 |
| 이사추적→거절 | 0.6 |
| 강력거절 | 11.8 |
| 기타 | 4.3 |
| 전 체 | 24.5 |

(3) 지역별 유형별 대체가구 수

각각의 대체사유별로 지역을 구분하여 보면 아래의 표와 같은데, 각각의 사유들을 합하면 전체에 해당하는 비율이 된다. 서울의 경우에 강력 거절이 가장 높은 비중을 차지하고 있는데 반하여, 울산의 경우에는 이사→추적 불가가 5.8%로 강력거절 5.0%에 비하여 높게 나타났다. 경남의 경우도 이사→추적 불가가 7.2%로 다른 사유에 비하여 가장 높게 나타났다.

(단위 : %)

| 지 역 | 주소 불명 | 이사 →추적 불가 | 이사 추적 →거절 | 강력 거절 | 기타 | 전체 |
|---|---|---|---|---|---|---|
| 서 울 | 1.0 | 7.3 | 1.0 | 14.8 | 6.2 | 30.2 |
| 부 산 | 4.3 | 4.7 | 0.2 | 8.2 | 1.9 | 19.4 |
| 대 구 | 0.6 | 8.8 | 0.6 | 14.4 | 0.9 | 25.3 |
| 인 천 | 0.7 | 3.7 | 1.0 | 11.2 | 6.8 | 23.4 |
| 광 주 | 0.0 | 5.0 | 0.6 | 10.6 | 0.0 | 16.3 |
| 대 전 | 1.9 | 3.2 | 0.0 | 11.0 | 1.9 | 18.1 |
| 울 산 | 3.3 | 5.0 | 0.0 | 8.3 | 0.0 | 16.7 |
| 경 기 | 1.9 | 4.6 | 0.6 | 12.8 | 7.9 | 27.7 |
| 강 원 | 4.6 | 1.5 | 0.0 | 14.6 | 6.9 | 27.7 |
| 충 청 | 1.7 | 3.8 | 0.0 | 8.9 | 1.7 | 16.2 |
| 전 라 | 2.5 | 3.7 | 0.9 | 11.7 | 3.7 | 22.5 |
| 경 북 | 0.0 | 14.8 | 0.7 | 8.1 | 0.4 | 24.1 |
| 경 남 | 1.7 | 7.6 | 0.7 | 5.2 | 1.7 | 16.9 |
| 전 국 | 1.7 | 6.1 | 0.6 | 11.8 | 4.3 | 24.5 |

4. 이사 가구

(1) 이사 가구 분포와 성공률

조사대상 가구 중 이사 가구는 전체의 11.3%인 563가구였으며, 이 중 40%인 226가구를 조사 성공하였다.

(단위 : %)

| 지 역 | 파악된 총 이사가구 | 이사 추적 불가 | 이사 추적 →거절 | 이사 추적 →성공 |
|---|---|---|---|---|
| 서 울 | 10.9 | 7.3 | 1.0 | 2.6 |
| 부 산 | 13.2 | 4.7 | 0.2 | 8.2 |
| 대 구 | 13.8 | 8.8 | 0.6 | 4.4 |
| 인 천 | 10.5 | 3.7 | 1.0 | 5.8 |
| 광 주 | 12.5 | 5.0 | 0.6 | 6.9 |
| 대 전 | 13.5 | 3.2 | 0.0 | 10.3 |
| 울 산 | 11.7 | 5.0 | 0.0 | 6.7 |
| 경 기 | 8.7 | 4.6 | 0.6 | 3.5 |
| 강 원 | 5.4 | 1.5 | 0.0 | 3.8 |
| 충 청 | 7.2 | 3.8 | 0.0 | 3.4 |
| 전 라 | 6.8 | 3.7 | 0.9 | 2.2 |
| 경 북 | 17.4 | 14.8 | 0.7 | 1.9 |
| 경 남 | 18.3 | 7.6 | 0.7 | 10.0 |
| 전 국 | 11.3 | 6.1 | 0.6 | 4.5 |

# V. 표본의 가중치

## 1. 한국노동패널 1차년도 가중치 부여방법

(1) 기본적인 고려사항

한국노동패널 1차년도에서는 가구가중치(household weights), 개인가중치(individual weights), 기록가구원가중치(enumerated individual weights)[1], 횡단면 노동(cross-sectional labor weights) 가중치[2] 등을 고려할 수 있다. 또한 가중치의 구성요소를 선정할 때, 일반적으로 가중치를 구하는 경우에 사용되는 추출확률, 응답률, 사후층화 요인 중에서 어떠한 요인들을

1) BHPS 참조
2) SLID 참조

포함할 것인가의 문제가 중요하다.

(2) KLIPS 표본개요[3]

한국노동패널 1차년도에서는 2단 층화집락계통추출법을 사용하였다. 먼저 1단계에서는 1995년 10% 센서스 표본조사구 중에서 도시지역 조사구 19,025개를 먼저 선정한 다음 이 중에서 951개의 조사구를 계통추출하였다. 이에 따라 서울, 부산, 대구, 인천, 광주, 대전, 울산 등의 광역시에서는 총 534 조사구가 선정되었으며, 기타 지역의 도시지역에서 417 조사구가 선정되었다(<표 2> 참조)[4].

<표2> 한국노동패널 1차년도의 표본개요(단위:개, 가구, %)

| 지역 | 센서스10%조사구 | | KLIPS 조사구 | 조사구추출률 | 추출가구수 | 응답가구수 | 응답률 |
|---|---|---|---|---|---|---|---|
| | 조사구수 | 도시조사구수 | | | | | |
| 서울 | 5,186 | 5,186 | 227 | 4.37716930 | 1,603 | 1,362 | 85.0 |
| 부산 | 1,841 | 1,841 | 97 | 5.26887561 | 600 | 485 | 80.8 |
| 대구 | 1,212 | 1,212 | 64 | 5.28052805 | 452 | 320 | 70.8 |
| 인천 | 1,116 | 1,116 | 59 | 5.28673835 | 414 | 295 | 71.3 |
| 광주 | 607 | 607 | 32 | 5.27182867 | 197 | 160 | 81.2 |
| 대전 | 592 | 592 | 31 | 5.23648649 | 188 | 155 | 82.4 |
| 울산 | 456 | 456 | 24 | 5.26315789 | 150 | 120 | 80.0 |
| 경기 | 3,624 | 3,542 | 167 | 5.15114127 | 1,223 | 853 | 69.7 |
| 강원 | 725 | 504 | 26 | 5.15873016 | 200 | 130 | 65.0 |
| 충북 | 694 | 422 | 22 | 5.21327014 | 145 | 110 | 75.9 |
| 충남 | 965 | 488 | 25 | 5.12295082 | 152 | 125 | 82.2 |
| 전북 | 967 | 757 | 0 | 5.28401585 | 273 | 200 | 73.3 |
| 전남 | 996 | 487 | 25 | 5.13347023 | 164 | 125 | 76.2 |
| 경북 | 1,372 | 1,021 | 54 | 5.28893242 | 356 | 270 | 75.8 |
| 경남 | 1,422 | 1,094 | 58 | 5.30154534 | 359 | 290 | 80.8 |
| 전국 | 21,675 | 19,025 | 951 | 4.99868594 | 6,476 | 5,000 | 77.2 |

주 : 제주도 263개 조사구 제외

2단계에서는 추출된 조사구로부터 서울 지역은 6가구, 기타 지역은 5가구를 임의추출하였다.[5] 실제로 2단계에서 최초로 추출한 가구숫자는 목표표본 가구였으며, 이러한 목표표본 가

3) 보다 구체적인 사항은 방하남 외(1999), 『한국가구와 개인의 경제활동』, 한국노동연구원 참조.
4) 실제로는 이렇게 계통추출된 조사구 중에서 고용구조특별조사의 조사구와 가장 가까운 조사구를 matching하여 추출하였다.
5) 단, 경기도는 총 167개 조사구 중에 18개 조사구에서 6가구를, 나머지 149개 조사구에서 5가구를 추출하였다.

구를 달성하기 위하여 조사구별로 접촉한 가구수는 5가구로부터 20가구까지 지역별로 매우 다양하게 나타나고 있다(<표 3> 참조). 한편, 추출된 가구의 경우에도 조사가 어려운 경우 대체표본을 사용하였으며, 이러한 대체표본은 전체 조사가구의 24.5%인 1,226가구이다(<표 4>참조).

(3) 표본추출확률의 계산

이와 같은 표본추출방법 하에서 추출확률은 다음과 같이 계산할 수 있다. 먼저, 표본조사구 추출확률은 센서스 10% 표본조사구로 선정되고 다시 KLIPS의 표본조사구로 선정될 확률이다. 이때 센서스 10% 표본조사구의 추출확률은 모든 지역에서 10%로 동일하다고 가정한다. 센서스 10% 표본조사구 내에서 최종적으로 표본조사구를 선정하는 경우에는 10% 표본조사구에서 선정한 조사구와 고용구조특별조사에서 선정한 조사구를 연결(matching)시키는 과정을 감안하여야 하는데 실제로 이 확률을 계산하는 것은 매우 어렵다. 이와 같은 점을 감안하여 표본조사구 추출확률은 서울 등 광역시의 경우에는 0.1*(KLIPS 표본조사구의 수/10% 센서스조사구의 수), 기타 지역의 경우에는 0.1*(도시조사구의 수/10% 센서스구의 수) *(KLIPS 표본조사구의 수/도시조사구의 수)로 계산하였다.

<표 3> 총접촉가구수의 조사구별 분포

| 구분 | 총가구수 | 조사구 | 서울 | 기타지역 |
|---|---|---|---|---|
| 5 | 1,205 | 241 | - | 241 |
| 6 | 1,077 | 209 | 31 | 178 |
| 7 | 825 | 155 | 44 | 111 |
| 8 | 565 | 107 | 29 | 78 |
| 9 | 504 | 92 | 40 | 52 |
| 10 | 311 | 56 | 28 | 28 |
| 11 | 115 | 21 | 10 | 11 |
| 12 | 151 | 27 | 14 | 13 |
| 13 | 96 | 17 | 11 | 6 |
| 14 | 53 | 9 | 8 | 1 |
| 15 | 26 | 5 | 1 | 4 |
| 16 | 36 | 6 | 5 | 1 |
| 18 | 6 | 1 | 1 | 0 |
| 19 | 18 | 3 | 3 | 0 |
| 20 | 12 | 2 | 2 | 0 |
| 계 | 5,000 | 951 | 227 | 724 |

<표 4> 한국노동패널 1차년도의 대체가구(단위:가구, %)

| 지역 | KLIPS총조사가구 | | | | 대체가구 | | |
|---|---|---|---|---|---|---|---|
| | 조사가구수 | 비중 | 조사가구원수 | 비중 | 가구수 | 지역별조사가구수 대비 비중 | 전체대체가구수 대비비중 |
| 서울 | 1,362 | 27.2 | 3,539 | 26.6 | 411 | **30.2** | 33.5 |
| 부산 | 485 | 9.7 | 1,330 | 10.0 | 94 | 19.4 | 7.7 |
| 대구 | 320 | 6.4 | 878 | 6.6 | 81 | **25.3** | 6.6 |
| 인천 | 295 | 5.9 | 825 | 6.2 | 69 | 23.4 | 5.6 |
| 광주 | 160 | 3.2 | 439 | 3.3 | 26 | 16.3 | 2.1 |
| 대전 | 155 | 3.1 | 426 | 3.2 | 28 | 18.1 | 2.3 |
| 울산 | 120 | 2.4 | 333 | 2.5 | 20 | 16.7 | 1.6 |
| 경기 | 853 | 17.1 | 2,155 | 16.2 | 236 | **27.7** | 19.2 |
| 강원 | 130 | 2.6 | 333 | 2.5 | 36 | **27.7** | 2.9 |
| 충청 | 235 | 4.7 | 678 | 5.1 | 38 | 16.2 | 3.1 |
| 전라 | 325 | 6.5 | 878 | 6.6 | 73 | 22.5 | 6.0 |
| 경북 | 270 | 5.4 | 692 | 5.2 | 65 | 24.1 | 5.3 |
| 경남 | 290 | 5.8 | 811 | 6.1 | 49 | 16.9 | 4.0 |
| 전국 | 5,000 | 100 | 13,303 | 100 | 1,226 | 24.5 | 100 |

최종 조사가구 선정확률을 엄밀하게 계산하려면 특정조사구에서 특정가구가 선정될 확률은 동일한 조사구 내에서 특정가구를 접촉하기 전에 미리 접촉 시도한 가구의 응답여부 또는 응답가구수에 따른 조건부확률로 계산되어야 한다. 예를 들어 서울지역의 어떤 조사구에서 7가구가 추출되었다면 6번째 가구까지는 동일한 조사구 내에서 임의적으로 선정되었기 때문에 동일한 추출확률을 가지고 있지만, 일곱 번째 접촉시도한 가구가 추출될 확률은 이전까지 접촉시도한 6가구 중에서 정확하게 1가구가 응답거절인 경우에 7번째 가구가 선정될 조건부 확률이다. 따라서 서울의 경우 6가구를 초과하는 가구의 추출확률, 기타 지역의 경우 5가구를 초과할 경우의 가구추출확률은 가구별로 상이하게 된다.

그러나 응답거절가구에 대한 정보가 전혀 없기 때문에 이러한 조건부확률을 구하기는 현실적으로 불가능하다. 이와 같은 점을 감안하여 특정한 조사구내에서 가구가 추출될 확률은 동일한 조사구내에서 접촉시도한 모든 가구의 추출확률(조사구내의 총접촉시도가구/조사구내의 총가구수)로 계산한다. 최종적으로 가구의 추출확률은 조사구추출확률과 조사구내에서의 가구추출확률의 곱으로 계산한다.

(4) 응답률과 면접확률의 계산

일반적으로 응답률조정을 위해서는 무응답가구에 대한 특성파악이 전제되어야 한다. 그러

나 한국노동패널 1차년도의 경우에는 무응답가구의 지역(광역시 또는 기타 道의 대도시) 이외에는 알려진 특성이 없다. 이에 따라 응답률조정은 지역요인만을 고려한다(지역별 응답률은 <표 2>참조). 최종적으로 추출확률과 응답률을 감안하면 특정한 가구의 가중치는 가구추출확률과 가구가 속한 지역의 응답률의 곱의 역수로 계산한다.

(5) 사후층화

사후층화는 가구단위에서의 사후층화와 개인단위에서의 사후층화를 고려할 수 있다. 그러나 가구단위의 정보는 1995년 센서스에 의한 자료가 가장 최근 자료이라는 점을 고려할 때, 이를 이용하여 사후층화를 하는 것은 시의성이 떨어진다. 뿐만 아니라, 개인단위에서 사용할 수 있는 자료는 통계청의 인구추계자료가 있으나 이는 지역별, 연령별 세분화가 되어 있지 않기 때문에(서울 및 광역시 지역의 경우에는 연령별, 성별 인구분포가 있으나 기타지역의 경우에는 전체 지역에 대한 연령별, 성별추계는 있으나 시지역에 대한 추계는 가용하지 않았음), 개인단위의 사후층화를 고려하는 것도 어렵다. 기타 경제활동인구조사나 고용구조특별조사의 결과를 이용하여 사후층화를 고려할 수도 있으나 두 조사 모두 표본조사이므로 이들 자료를 이용하여 사후층화하는 것은 사후층화의 의미가 퇴색된다. 이와 같은 점을 고려하여 사후층화는 하지 않는다.

이에 따라 결과적으로는 동일한 가구내의 가구원은 모두 동일한 가중치를 부여받게 된다.

## 2. 가중치 부여 결과

(1) 가구용 데이터

가중치는 추출확률만을 계산한 w_p와 응답률까지 고려한 w_p_r 등 2가지로 제공된다. w_p는 평균 1,583.859이고, 최소 468.339에서 최대 4,554.800의 값을 가지며, w_p_r은 평균 2,063.240이고, 최소 551.210에서 최대 5,540.629를 가진다(<표 5>참조).

<표 5> 가중치의 요약통계량 : 가구용데이터

| 구분 | 평균 | 표준편차 | 최소 | 최대 |
|---|---|---|---|---|
| w_p | 1,583.859 | 670.326 | 468.339 | 4,554.800 |
| w_p_r | 2,063.240 | 902.022 | 551.210 | 5,540.629 |

<표 6> 지역별 가중치의 요약통계량 : 가구용데이터

| 구분 | w_p | | w_p_r | |
|---|---|---|---|---|
| | 평균 | 표준편차 | 평균 | 표준편차 |
| 서울 | 1,107.841 | 299.284 | 1,303.869 | 352.2412 |
| 부산 | 1,380.567 | 246.221 | 1,707.918 | 304.6045 |
| 대구 | 1,212.200 | 314.515 | 1,712.233 | 444.2538 |
| 인천 | 1,210.954 | 311.691 | 1,699.440 | 437.4251 |
| 광주 | 1,590.485 | 302.263 | 1,958.285 | 372.1622 |
| 대전 | 1,662.084 | 274.836 | 2,015.947 | 333.3503 |
| 울산 | 1,576.443 | 316.845 | 1,970.554 | 396.0570 |
| 경기 | 1,605.212 | 394.332 | 2,301.494 | 565.3796 |
| 강원 | 2,037.860 | 596.061 | 3,135.169 | 917.0183 |
| 충북 | 2,630.375 | 649.124 | 3,467.313 | 855.6639 |
| 충남 | 3,471.469 | 758.718 | 4,221.306 | 922.6021 |
| 전북 | 1,888.998 | 440.517 | 2,578.483 | 601.3058 |
| 전남 | 3,228.545 | 715.389 | 4,235.851 | 938.5904 |
| 경북 | 2,063.790 | 475.531 | 2,721.145 | 626.9964 |
| 경남 | 2,063.430 | 393.328 | 2,554.384 | 486.9132 |
| 전체 | 1,583.858 | 670.326 | 2,063.239 | 902.022 |

<표 6>에 의하면 서울의 경우가 가중치가 가장 작고 <표 7>에 의하면 서울지역의 표본의 구성비가 가중치를 사용하지 않는 경우에 비해 매우 낮아짐을 알 수 있다. 즉, 가중치를 사용하지 않은 경우의 서울가구의 구성비가 27.06%인데 반해, 가중치로 조정한 서울의 표본가구의 비중은 w_p를 사용한 경우 전체의 19.05%, w_p_r을 사용한 경우 전체의 17.26%에 불과하게 된다. 이는 다른 지역의 조사구에서는 5가구를 선정한 반면, 서울지역에서는 6가구를 선정하였기 때문이며, 또한 서울지역의 응답률이 전국에서 가장 높게 나타났기 때문이다.

<표 8>을 보면 가중치를 사용하지 않는 경우의 가구 월평균 근로소득은 131.9만원이며 w_p를 사용한 경우 129.2만원, w_p_r을 사용한 경우 129.0만원으로 나타나 큰 차이를 보이지는 않는다. 가구 월평균 지출액의 경우에는 가중치를 사용하지 않는 경우 101.2만원이나 w_p를 사용하는 경우 97.6만원, w_p_r을 사용하는 경우 97.4만원으로 나타나서 상대적으로 큰 차이를 보이고 있다.

<표 7> 가중치와 지역별 표본수 : 가구용데이터

| 구분 | No weight | | w_p | | w_p_r | |
|---|---|---|---|---|---|---|
| | 표본수 | 비중 | 표본수 | 비중 | 표본수 | 비중 |
| 서울 | 1,353 | 27.06 | 150,7824 | 19.05 | 1,779,998 | 17.26 |
| 부산 | 486 | 9.72 | 67,1728 | 8.48 | 831,618 | 8.06 |
| 대구 | 314 | 6.28 | 38,0100 | 4.80 | 536,545 | 5.20 |
| 인천 | 153 | 3.06 | 25,4859 | 3.22 | 309,182 | 3.00 |
| 광주 | 298 | 5.96 | 36,5023 | 4.61 | 511,244 | 4.96 |
| 대전 | 158 | 3.16 | 25,1703 | 3.18 | 309,975 | 3.01 |
| 울산 | 116 | 2.32 | 18,2657 | 2.31 | 228,625 | 2.22 |
| 경기 | 856 | 17.12 | 136,1049 | 17.19 | 1,946,509 | 18.87 |
| 강원 | 130 | 2.60 | 26,4855 | 3.35 | 407,485 | 3.95 |
| 충북 | 113 | 2.26 | 29,3114 | 3.70 | 386,349 | 3.75 |
| 충남 | 131 | 2.62 | 44,5170 | 5.62 | 543,342 | 5.27 |
| 전북 | 199 | 3.98 | 37,6297 | 4.75 | 513,036 | 4.97 |
| 전남 | 127 | 2.54 | 40,1390 | 5.07 | 526,786 | 5.11 |
| 경북 | 274 | 5.48 | 56,0358 | 7.08 | 738,693 | 7.16 |
| 경남 | 292 | 5.84 | 60,0832 | 7.59 | 744,350 | 7.22 |
| 전체 | 5,000 | 100.00 | 7,916,959 | 100.00 | 10,313,737 | 100.00 |

<표 8> 가중치와 소득관련변수

(단위: 가구, 만원)

| 구 분 | | 표본수 | 평균 | 표준편차 |
|---|---|---|---|---|
| 가구근로소득 (v01441) | No weight | 4,928 | 131.903 | 121.994 |
| | w_p | 7,827,311 | 129.208 | 120.233 |
| | w_p_r | 10,196,283 | 129.032 | 119.923 |
| 한달평균생활비 (v01454) | No weight | 4,986 | 101.208 | 71.339 |
| | w_p | 7,900,083 | 97.552 | 68.092 |
| | w_p_r | 10,290,581 | 97.421 | 67.774 |

<표 9>는 1998년 통계청의 도시가계연보에 나타난 도시근로자가구의 월평균 소득과 한국노동패널 1차년도에서의 2인 이상 도시근로자가구의 가구소득을 비교한 것이다[6]. KLIPS에

6) 통계청에서 사용하는 가구소득은 KLIPS에서 가구총소득과 대응되는 개념이다. 통계청의 가구소득은 경상소득과 비경상소득의 합으로 정의되며, 가구 경상소득은 KLIPS의 근로소득, 재산소득, 이전소득의 합과 같다고 볼 수 있으며, 통계청의 비경상소득은 KLIPS의 기타소득과 유사한 개념으로 볼 수 있다. 다만 통계청의 가구경상소득에는 KLIPS에서 가구총소득의 일부로 계산되는 부동산매

서 2인 이상 근로자가구의 월평균 가구총소득은 174.7만원으로 통계청 월평균 가구소득 213.3만원의 81.9%에 불과하다. KLIPS에서의 가구 경상소득은 172.7만원으로 통계청 경상소득의 86.6%에 해당된다. 근로소득의 경우에는 동 비율이 88.5%로 상대적으로 유사하며, 재산소득의 경우에도 매우 유사하다(두 조사에서 재산소득의 정의가 다름에 유의하여야 한다 <각주 14>참조). 그러나 이전소득의 경우에는 동 비율이 13.8%에 불과하여 KLIPS에서 상대적으로 적게 추정됨을 알 수 있다.[7] 한편, 이와 같은 결과는 가중치를 부여하여도 크게 변화하지 않는데, 가중치를 부여하지 않은 경우 가구총소득이 174.7만원인 반면 w_p_r을 사용한 가중평균치는 171.9만원으로 나타났다.

<표 9> 도시가계연보와 가구소득 비교

(단위:만원)

| 구분 | 소득 | 경상소득 | 근로소득 | 사업 및 부업소득 | 재산소득 | 이전소득 | 비경상소득 |
|---|---|---|---|---|---|---|---|
| 도시 가계연보 | 213.3 | 199.4 | 181.6 | 7.8 | 4.3 | 5.8 | 13.9 |
| | | | 189.4 | | | | |
| KLIPS1 | 174.7 | 172.7 | 167.7 | | 4.2 | 0.8 | 2.0 |
| KLIPS2 | 172.3 | 170.6 | 166.0 | | 3.9 | 0.7 | 1.7 |
| KLIPS3 | 171.9 | 170.3 | 165.7 | | 3.9 | 0.7 | 1.6 |

주: KLIPS1, KLIPS2, KLIPS3는 각각 가중치를 부여하지 않은 경우, w_p를 가중치로 사용한 경우, w_p_r을 가중치로 사용한 경우에 해당됨.
자료: 통계청, 도시가계연보(1998), KLIPS 1차 데이터

분석대상이 되는 표본가구수가 도시가계조사의 경우 3,098가구이며, KLIPS의 경우 2,018가구임을 감안하여 도시가계조사에서의 추정치가 상대적으로 母數에 유사할 가능성이 많다고 하면 KLIPS의 가구소득관련 추정치는 실제보다 적게 추정되었을 가능성이 있다. 또한 이러한 차이는 표본추출확률과 응답률로 보정되지 않는 소득격차라고 할 수 있겠다.

다음의 <표 10>는 자영업자가구의 소득현황을 보여주고 있는데, 자영업가구의 소득이 근로자가구의 소득에 비해 적은 것으로 나타나고 있다. 1998년의 자영업자 가구소득에 대한 공식적인 데이터는 없지만, 1996년을 기준으로 통계청이 실시한 96가구소비실태조사에 따르면 자영업가구의 연간 평균소득은 3,121.0만원으로 근로자가구의 연간 평균소득 2,455.6만원의

매차익, 주식의 매매차익 등은 제외된다는 점에서 포괄범위로 볼 때에는 KLIPS에서 재산소득이 통계청의 재산소득보다 적을 것으로 예상된다.

7) 강석훈(1999)은 통계청 96가구소비실태조사에서 추정된 이전소득이 한국은행 국민계정 이전소득의 20%내외임을 보인바 있다. 이와 같은 수준이 도시가계조사에서도 적용된다고 가정하면 KLIPS의 이전소득은 매우 과소추정되었다고 보여진다.

127.1%에 달하는 것으로 나타나고 있다. 한국노동패널 1차년도에서의 동 비율은 가구소득을 기준으로 93.5%에 불과하여 오히려 자영업자의 소득이 근로자가구에 비해 적은 것으로 나타났다. 1인가구를 제외한 자영업자의 소득은 근로자가구(1인가구 제외)의 96.5%로 나타나 역시 근로자가구보다 적은 수준임을 알 수 있다. 가중치를 부여하는 경우 자영업자의 소득은 근로자가가구에 비해 상대적으로 큰 차이를 보이고 있다. 전가구를 대상으로 할 때 자영업자의 월평균소득은 163.4만원이었으며, w_p를 가중치로 사용한 경우 156.5만원, w_p_r을 가중치로 사용한 경우 156.4만원으로 나타나 가중평균소득은 감소하는 것으로 나타났다.

이와 같은 결과는 1998년 IMF위기에서 자영업자의 소득감소폭이 컸다고 해석할 수도 있고, 또는 KLIPS에서 자영업자의 소득이 근로가구의 소득보다 상대적으로 많이 축소보고 되었다고도 볼 수 있으며 향후 더욱 연구하여야 할 과제이다.

<표 10> 자영업자의 월별소득

(단위:만원)

| 구분 | | 가구소득 | 근로소득 | 재산소득 | 이전소득 | 기터소득 |
|---|---|---|---|---|---|---|
| 전가구 | KLIPS1 | 163.4 | 153.2 | 6.0 | 1.7 | 2.5 |
| | KLIPS2 | 156.5 | 146.9 | 5.4 | 1.9 | 2.4 |
| | KLIPS3 | 156.4 | 146.6 | 5.4 | 1.9 | 2.5 |
| 1인 가구제외 | KLIPS1 | 168.5 | 157.9 | 6.3 | 1.7 | 2.6 |
| | KLIPS2 | 161.5 | 151.6 | 5.7 | 1.7 | 2.5 |
| | KLIPS3 | 161.3 | 151.3 | 5.6 | 1.8 | 2.6 |

주: KLIPS1, KLIPS2, KLIPS3는 각각 가중치를 사용하지 않은 경우, w_p를 가중치로 사용한 경우, w_p_r을 가중치로 사용한 경우임.

(2) 개인용 데이터

개인용 가중치의 경우 w_p는 평균 1,570.617, 최소 468.339, 최대 4,223.040으로 나타났으며, w_p_r의 경우 평균 2,154.610, 최소 551.209, 최대 7,259.752로 나타났다(<표 11> 참조). 지역별로 볼 때 서울이 평균 1,104.661로 가장 작게 나타났으며, 전남의 경우 3,175.461로 가장 크게 나타났다. 따라서 서울이 차지하는 비중이 가중치를 사용하지 않는 경우 27.27%였으나 w_p를 사용하는 경우 19.18%, w_p_r을 사용하는 경우 16.45%에 불과하였다(<표 12>, <표 13> 참조).

<표 14>는 임금근로자의 월평균소득을 나타내고 있는데 가중치를 사용하지 않는 경우 111.5만원이었으며, w_p를 사용하는 경우 110.2만원, w_p_r을 사용하는 경우 109.1만원으로 큰 차이는 없음을 보여주고 있다[8].

8) 한편, 통계청에 의하면 1998년 상용근로자의 월평균임금은 142.7만원이며, KLIPS에서 정규직의 월

<표 15>에는 지난 1주간 주요활동(KLIPS 코드북의 변수명 p01568)에 대한 분포를 나타내고 있는데 일하였음과 일시휴직, 그리고 구직활동의 경우를 경제활동인구로 간주하고, 구직활동을 실업자로 간주하는 경우 가중치를 사용하지 않으면 실업률은 12.79%, w_p를 사용하면 실업자와 실업률이 각각 140.6만명, 12.11%, w_p_r을 사용하면 실업자와 실업률이 각각 191.0만명, 12.00%에 이르는 것으로 추정된다.

<표 11> 가중치의 요약통계량 : 개인용데이터

| 구분 | 평균 | 표준편차 | 최소 | 최대 |
|---|---|---|---|---|
| w_p | 1,570.617 | 640.561 | 468.339 | 4,223.040 |
| w_p_r | 2,154.610 | 1,085.899 | 551.209 | 7,259.752 |

<표 12> 지역별 가중치의 요약통계량 : 개인용데이터

| 구분 | w_p | | w_p_r | |
|---|---|---|---|---|
| | 평균 | 표준편차 | 평균 | 표준편차 |
| 서울 | 1,104.661 | 297.045 | 1,300.126 | 349.606 |
| 부산 | 1,382.283 | 242.024 | 1,710.041 | 299.412 |
| 대구 | 1,209.845 | 309.907 | 1,708.906 | 437.744 |
| 인천 | 1,221.348 | 303.656 | 1,714.028 | 426.148 |
| 광주 | 1,592.822 | 304.829 | 1,961.163 | 375.321 |
| 대전 | 1,665.395 | 277.358 | 2,019.963 | 336.408 |
| 울산 | 1,590.899 | 306.859 | 1,988.623 | 383.574 |
| 경기 | 1,606.983 | 393.142 | 2,304.034 | 563.673 |
| 강원 | 2,043.757 | 595.422 | 3,144.242 | 916.034 |
| 충북 | 2,584.985 | 676.950 | 3,407.480 | 892.343 |
| 충남 | 3,158.410 | 654.618 | 3,840.627 | 796.015 |
| 전북 | 1,871.266 | 447.325 | 5,108.556 | 1221.19 |
| 전남 | 3,175.461 | 722.004 | 4,166.205 | 947.269 |
| 경북 | 2,083.481 | 471.708 | 2,747.108 | 621.957 |
| 경남 | 2,081.916 | 389.556 | 2,577.268 | 482.243 |
| 전체 | 1,570.617 | 640.561 | 2,154.610 | 1,085.898 |

평균임금은 123.7만원, w_p_r을 가중치로 사용하는 경우는 121.6만원으로 나타났다. 통계청의 상용근로자와 KLIPS의 정규직근로자가 정확하게 대응되는 개념은 아니라는 점을 감안하더라도 가구소득에서와 마찬가지로 KILPS에서의 추정치가 통계청의 추정치보다 적게 나타나고 있다.

<표 13> 가중치와 지역별 표본수 : 개인용데이터

| 구분 | No weight | | w_p | | w_p_r | |
|---|---|---|---|---|---|---|
| | 표본수 | 비중 | 표본수 | 비중 | 표본수 | 비중 |
| 서울 | 3,746 | 27.27 | 4,136,416 | 19.18 | 4,868,600 | 16.45 |
| 부산 | 1,325 | 9.64 | 1,830,851 | 8.49 | 2,265,187 | 7.65 |
| 대구 | 942 | 6.86 | 1,139,238 | 5.28 | 1,609,250 | 5.44 |
| 인천 | 835 | 6.08 | 1,019,455 | 4.73 | 1,430,680 | 4.83 |
| 광주 | 448 | 3.26 | 713,275 | 3.31 | 878,415 | 2.97 |
| 대전 | 441 | 3.21 | 734,258 | 3.40 | 890,589 | 3.01 |
| 울산 | 338 | 2.46 | 537,534 | 2.49 | 672,062 | 2.27 |
| 경기 | 2,254 | 16.41 | 3,621,279 | 16.79 | 5,192,113 | 17.54 |
| 강원 | 350 | 2.55 | 715,134 | 3.32 | 1,100,258 | 3.72 |
| 충북 | 307 | 2.23 | 793,433 | 3.68 | 1,045,946 | 3.53 |
| 충남 | 331 | 2.41 | 1,045,300 | 4.85 | 1,271,104 | 4.30 |
| 전북 | 584 | 4.25 | 1,092,505 | 5.06 | 2,982,980 | 10.08 |
| 전남 | 336 | 2.45 | 1,066,822 | 4.95 | 1,399,675 | 4.73 |
| 경북 | 698 | 5.08 | 1,453,878 | 6.74 | 1,917,305 | 6.48 |
| 경남 | 803 | 5.85 | 1,671,358 | 7.75 | 2,069,198 | 6.99 |
| 전체 | 13,738 | 100.00 | 21,570,736 | 100.00 | 29,593,362 | 100.00 |

<표 14> 가중치와 소득관련변수

(단위: 명, 만원)

| 구 분 | | 표본수 | 평균 | 표준편차 |
|---|---|---|---|---|
| 월평균임금 (p01091) | No weight | 3,974 | 111.539 | 67.029 |
| | w_p | 6,130,457 | 110.153 | 65.373 |
| | w_p_r | 8,391,269 | 109.128 | 65.089 |

<표 15> 지난 1주간 주요활동(p01568)

| 구분 | No weight | | w_p | | w_p_r | |
|---|---|---|---|---|---|---|
| | 표본수 | 비중 | 표본수 | 비중 | 표본수 | 비중 |
| 일하였음 | 6281 | 45.72 | 10040951 | 46.55 | 13737258 | 46.42 |
| 임시휴직 | 105 | 0.76 | 169522 | 0.79 | 265044 | 0.90 |
| 구직활동 | 937 | 6.82 | 1406028 | 6.52 | 1909726 | 6.45 |
| 가사/육아 | 2280 | 16.60 | 3482690 | 16.15 | 4746783 | 16.04 |
| 학업 | 2009 | 14.62 | 3118416 | 14.46 | 4277423 | 14.45 |
| 연로 | 1058 | 7.70 | 1693482 | 7.85 | 2337623 | 7.90 |
| 심신장애 | 307 | 2.23 | 520468 | 2.41 | 727532 | 2.46 |
| 기타 | 761 | 5.54 | 1139179 | 5.28 | 1591973 | 5.38 |
| 전체 | 13738 | 100.00 | 21570736 | 100.00 | 29593362 | 100.00 |

## 3. 제한점

(1) 가중치부여 방법론의 문제

본고에서 고려한 가중치 부여방법은 다음과 같은 사항을 고려하지 못하고 있다.

첫째, 이사가구 문제이다. 실제 한국노동패널 1차년도의 면접과정에서는 추출된 주소에서

<표 16> 이사가구의 분포와 추적

(단위:%)

| 구분 | 파악된 총 이사가구 | 이사-추적불가 | 이사추적-거절 | 이사추적-성공 |
|---|---|---|---|---|
| 서울 | 10.9 | 7.3 | 1.0 | 2.9 |
| 부산 | 13.2 | 4.7 | 0.2 | 8.2 |
| 대구 | 13.8 | 8.8 | 0.6 | 4.4 |
| 인천 | 10.5 | 3.7 | 1.0 | 5.8 |
| 광주 | 12.5 | 5.0 | 0.6 | 6.9 |
| 대전 | 13.5 | 3.2 | 0.0 | 10.3 |
| 울산 | 11.7 | 5.0 | 0.0 | 6.7 |
| 경기 | 8.7 | 4.6 | 0.6 | 3.5 |
| 강원 | 5.4 | 1.5 | 0.0 | 3.8 |
| 충청 | 7.2 | 3.8 | 0.0 | 3.4 |
| 전라 | 6.8 | 3.7 | 0.9 | 2.2 |
| 경북 | 17.4 | 14.8 | 0.7 | 1.9 |
| 경남 | 18.3 | 7.6 | 0.7 | 10.0 |
| 전국 | 11.3 | 6.1 | 0.6 | 4.5 |

의 가구가 이사를 한 경우에는 이 가구를 추적하는 방법을 사용하였다. 조사대상가구 중 이사가구는 전체의 11.3%인 563가구였으며, 이중 4.5%인 226가구를 추적하여 조사하였다. 그러나 이사하여 추적조사한 가구의 추출확률은 계산불가능하기 때문에 이러한 요인을 가중치에 반영하지 못하였다[9].

둘째, 대체가구 문제이다. 본문에서도 지적하였듯이 조사가구의 24.5%인 1,226가구가 대체가구에 의해 설문서가 작성되었는데 이러한 대체가구의 추출확률을 계산하는 것은 불가능하다.

마지막으로, 응답률 계산 문제이다. 본 연구에서의 가중치는 광역시나 개별 도내에서의 응답률을 기준으로 응답률을 계산하였으나 동일 지역내에서도 조사구별로 異質性이 클 수 있다는 점을 감안할 때 보다 정교한 응답률 계산방법을 고려할 필요가 있다. 현재 주어진 정보로는 보다 세분화된 응답률을 산출할 수 없으나 향후 개별 조사구에 대한 추가적인 정보가 가용할 때는 응답률 세분화 작업이 필요할 것으로 보이며, 사후층화의 문제도 동일한 맥락에서 접근될 것이다.

(2) 가중치부여 결과의 해석

현재 작성된 가중치를 사용하는 경우에 지역별로 하나의 가구나 개인이 나타내는 대표성은 크게 차이나고 있으나, 중요 변수인 가구소득이나 임금수준, 실업률은 가중치를 사용하지 않은 경우와 큰 차이를 보이지 않고 있다.

그러나 한국노동패널 1차년도의 분석결과를 공식적으로 사용되는 통계청의 추계와 비교하면 대체로 KLIPS가 통계청의 추정결과에 비해 실업률은 상대적으로 높고 소득은 상대적으로 낮은 수준을 보여주고 있다.

본 연구의 가중치부여결과에 의하면 표본추출확률과 응답률을 보정하여도 이러한 차이가 크게 축소되지는 않고 있음을 알 수 있다. 가중치부여과정의 문제점을 제외하는 경우 이와 같은 차이는 다음과 같은 원인에 기인할 수 있다.

먼저, KLIPS 표본이 지역별로 큰 차이가 없이 상대적으로 실업률이 높고 경제활동참가율이 낮고 따라서 가구소득이 적은 가구를 동일 유형의 모집단의 가구비율에 비해 상대적으로 과다포함하고 있을 가능성이 있다는 점을 지적할 수 있다.

또한, KLIPS의 면접과정에서 어떠한 이유로든지 표본가구들이 소득을 상대적으로 축소보고 했을 가능성을 생각해 볼 수 있다[10]. KLIPS 1차 웨이브의 면접시기가 IMF경제위기가

9) 원천적으로는 이사가구를 추적하여 조사하는 방법이 거주지표본을 이용한 경우에 적합한 방법인가에 대한 논란의 여지가 있다.

10) KLIPS에서는 월평균근로소득이 2,000만원에서 top coding되었다. 이러한 top coding문제로 인하여 초고소득층이 분석결과에서 누락되었을 가능성도 있다.

최악의 국면을 보이고 있었던 1998년 하반기에 이루어졌다는 점을 감안할 때 당시의 사회분위기 등으로 인하여 응답자들이 실제 소득보다 적게 보고하였을 가능성도 고려할 수 있다. 물론, 실제로 모집단의 값들이 통계청의 추정치보다는 한국노동패널 1차년도의 추정치와 유사할 가능성도 배제할 수 없다.

## Ⅵ. 자료의 구성 및 내용

제1차 패널조사는 앞으로의 조사를 위한 기초 연도(base year) 조사이므로 현재의 경제활동상태 파악과 함께 개인의 배경 정보와 과거 직업 경력(job history)을 중심으로 구성하였다.

자료의 구성은 크게 가구용 자료와 개인용 자료로 나뉘어져 있다. 가구용 자료에는 가구원 개개인에 관련된 성, 가구주와의 관계, 학력과 지난 1주간 주요 활동 등이 있으며 그 외 주거 관련 내용과 초등학교 3학년 이하 자녀의 보육, 그리고 가구의 소득, 소비 관련 내용으로 구성된다. 개인용 자료는 15세 이상의 가구원 개인에 대해 조사한 내용으로, 가구원의 경제활동상태에 따라 임금근로자, 비임금근로자, 미취업자에 적합한 내용들로 구성되었다. 자료의 내용을 간략하게 정리하면 다음과 같다(자세한 내용은 코드북 참조).

### 1. 가구용 자료

1. 가구원의 인적사항
    - 성별
    - 가구주와의 관계
    - 동거여부와 이유
    - 생년월일
    - 학력
    - 지난 1주간 경제활동상태
2. 자녀 보육 관련 질문
    - 초등학교 3학년 이하의 자녀 유무
    - 보육시설, 교육기관 이용여부와 이용실태
    - 탁아모 이용여부와 실태
3. 주거 상태
    - 주택의 자가, 전·월세 여부

- 주택의 종류
- 주택의 평수
- 주택의 싯가, 전세 보증금, 월세
- 거주시작시기

4. 가구의 소득과 소비
   - 사회보험 수혜자 여부와 액수
   - 근로, 금융, 부동산, 이전, 기타 소득 여부와 액수
   - 월평균 저축액
   - 부채 여부와 액수
   - 월평균 생활비
   - 경제적 어려움 여부와 내용

## 2. 개인용 자료

### (1) 임금근로자 관련 자료

1. 일하는 곳(직장) 관련 사항
   - 사업체 이름과 업종
   - 하고 있는 일의 종류(직종)
   - 직위, 직명
   - 현재의 일 시작시기
   - 현직장의 종류
   - 사업체 규모
   - 고용형태(정규직, 비정규직 여부)
   - 근로계약기간 유무와 계약 기간
   - 근무시간 형태(시간제, 전일제)
   - 종사상 지위(상용, 임시, 일용)
2. 근로시간
   - 근로시간의 규칙성
   - 일자리의 정기 휴일
   - 교대제 여부와 종류
   - 정규근무시간 유무와 정규근무시간
   - (정규근무시간이 없는 경우)평소 근무시간
   - 초과근무시간 유무와 초과근무시간, 초과근무수당

- 지난주 실제 근무 시간

3. 임금지급방식
   - 임금 지급 방식
   - 임금 결정 방식과 액수
   - 성과급제 유무와 적용 성과급제 종류
   - 월평균 총임금(세금공제 후)과 내용(기본급, 상여금 등)
4. 부업 및 아르바이트
   - 부업 및 아르바이트 유무
   - 부업 및 아르바이트를 하는 이유
   - 사업체 이름, 업종
   - 직종
   - 근로 시간
   - 시작 시기
   - 월평균 수입
5. 이전 직장
   - 이전 직장의 사업체명, 업종
   - 주로 하던 일
   - 근로 기간
   - 일자리에서의 지위
   - 일자리의 규모
   - 일자리를 그만둔 이유
   - 그만둘 당시 월평균 임금
6. 앞으로의 계획
   - 현재 일자리 지속 여부
   - 희망하는 일자리의 종류, 고용형태, 근무시간
   - 구직 여부와 구직 방법
   - 구직시 어려움과 구직 기간
7. 기타
   - 노동조합 유무와 가입 여부
   - 직업만족도
   - 현재의 직장 취업 정보와 구직 방법
   - 생활비 부담 정도
   - 1997년 경제위기 이후 임금과 근무시간 변화 정도

**(2) 비임금근로자 관련 자료**

1. 사업체 관련
   - 사업체명과 업종
   - 주로 하는 일(직종)
   - 사업 시작 시기
   - 고용된 종업원 수
   - 무급가족 종사자 수
   - 사업의 대표자
   - 1년 매출액
   - 월평균 수입
2. 근로 시간
   - 근로 시간의 규칙성
   - 정기 휴일
   - 평균 근로 시간과 지난주 실제 근로 시간
3. 부업 및 아르바이트
   - 부업 및 아르바이트 유무
   - 부업 및 아르바이트를 하는 이유
   - 사업체 이름, 업종
   - 직종
   - 근로 시간
   - 시작 시기
   - 월평균 수입
4. 이전 직장
   - 이전 직장의 사업체명, 업종
   - 주로 하던 일
   - 근로 기간
   - 일자리에서의 지위
   - 일자리의 규모
   - 일자리를 그만둔 이유
   - 그만둘 당시 월평균 임금
5. 앞으로의 계획
   - 현재 일자리 지속 여부
   - 희망하는 일자리의 종류, 고용형태, 근무시간
   - 구직 여부와 구직 방법

• 구직시 어려움과 구직 기간

6. 기타

• 직업만족도
• 현사업 시작시 어려운 점과 현사업을 택한 이유
• 생활비 부담 정도
• 1997년 경제위기 이후 매출의 변화 정도

**(3) 미취업자 관련 자료**

1. 구직상태

• 지난주 구직 여부
• 지난 1개월간 구직 여부
• 지난주 취업 가능성 여부

2. 구직자의 경우

• 구직 이유
• 희망 일자리의 업종, 직종, 고용형태, 근무시간
• 희망 수입
• 구직 방법
• 구직시 어려움
• 구직 기간
• 구직하였지만 일자리를 제의받지 못했던 이유

3. 비경제활동인구의 경우

• 현재 근로 가능성 여부
• 구직 이유
• 일자리를 구하지 않는 이유
• 일자리가 있어도 일하지 않는 이유
• 희망 일자리의 업종, 직종, 고용형태, 근무시간
• 희망 수입

4. 이전 직장

• 이전 직장의 사업체명, 업종
• 주로 하던 일
• 사업체의 종류와 규모
• 종사상 지위와 근로시간
• 근로계약여부와 계약기간
• 근로 기간

- 일자리에서의 지위
- 그만둘 당시 월평균 임금
- 퇴직금 수급 여부와 액수
- 일자리를 그만둔 이유

5. 기타
   - 가구의 생계 유지 방법

**(4) 개인 공통 자료**

1. 직업훈련
   - 직업훈련 수료 여부
   - 직업훈련의 종류
   - 직업훈련을 받은 장소
   - 직업훈련 교육 기간과 시간
   - 직업훈련 비용 부담자
   - 새로운 직업훈련 희망 여부와 받고 싶은 직업훈련
2. 정규 교육
   - 현재 정규학교 재학 여부와 졸업년도
   - 전문대 이상 교육 여부
   - 전문대 이상 학교명, 학과명, 재학 시기
   - 전문대 이상 휴학 이유와 휴학 시기
3. 군대 복무 여부와 기간
4. 아버지의 교육과 14세 무렵 직업
   - 아버지의 학력
   - 14세 무렵 아버지의 종사상 지위와 산업, 직업
5. 배경 질문
   - 출생지
   - 14세무렵 성장지
   - 생활만족도
   - 종교
   - 혼인상태와 혼인시기
   - 자녀 명수와 첫번째 자녀 출산 시기
   - 실업문제 해결을 위해 정부가 해야 할 일