

한국노동패널 표본의 대표성과 패널조사 표본 이탈자의 특성 연구

김대일 (서울대 경제학)

남재량 (한양대 경제학)

류근관 (서울대 경제학)

I. 서론

패널조사는 동일한 표본을 상이한 시점에서 반복하여 조사함으로써 횡단면 자료나 시계열 자료만으로는 얻을 수 없는 정보들을 제공한다는 점에서 그 필요성이 강조되어 왔다. 모든 통계조사에서 그렇듯이 패널조사 역시 수집된 자료가 모집단의 특성을 잘 대표할 수 있어야 한다. 특히 패널조사는 선정된 표본에 대한 조사가 1회에 그치지 않고 이후에도 계속된다는 점에서 표본의 대표성 문제뿐만 아니라, 확보된 대표성을 계속 유지하는 것이 매우 중요하다. 즉 최초에 추출되어 조사된 표본이 대표성을 유지한다고 하더라도 동일한 가구 또는 개인에 대한 조사가 시간에 걸쳐 계속됨에 따라 응답거부, 이사, 사망 등의 이유로 표본 이탈이 발생하므로 표본의 대표성을 유지하는 문제가 대두된다.

1998년부터 실시된 한국노동연구원의 「한국노동패널(KLIPS)」이 성공적인 패널조사로 정착되기 위해서는 표본의 대표성 확보 및 표본의 이탈을 최소화하기 위한 노력이 무엇보다도 중요하다. 따라서 본 연구는 다음 두 가지를 연구의 목적으로 한다. 첫째, 한국노동연구원에서 실시한 패널조사 결과를 기존의 통계자료와 비교 분석함으로써 한국노동패널(KLIPS)에서 선정한 표본이 과연 어느 정도의 대표성을 확보하고 있는지를 평가한다. 둘째, 기존에 구축된 타 패널자료를 사용하여 최초 선정된 표본 가운데 이후의 거듭된 조사에서 이탈하는 표본 이탈자들의 특성이 무엇인지를 분석한다. 이를 통해 패널조사에서 가장 중요한 문제인 표본 이탈률(attrition rate)을 낮추어 KLIPS가 성공적인 패널조사로 정착되기 위해 어떤 노력이 필요한 지에 대한 시사점을 얻는다.

본 연구의 구성은 다음과 같다. 제2장은 KLIPS에서 선정하여 조사한 표본의 대표성을 기존의 통계자료와 비교하여 분석한다. 제3장은 국내에서 구축된 기존의 패널조사 자료를 사용하여 최초 선정된 표본에서 이탈하는 사람들의 특성에 대해 분석한다. 마지막으로 제4장은 이상의 분석을 요약하고 이를 바탕으로 KLIPS가 성공적인 패널조사로 자리잡기 위해 필요한 노력들에 대해 시사점을 제시한다.

II. 한국노동패널 표본의 대표성

1. 개괄

한국 노동연구원 패널은 1차년도인 1998년, 5000 가구 13738명을 대상으로 가구 및 개인에 대한 자료를 조사하고 있다. 가구에 대해서는 가구원 구성 및 각종 소득 등을 조사하고 있고, 개인에 대해서는 과거 및 현재의 경제활동 상황, 그리고 소득에 대한 자료를 조사하였다. 한국 노동연구원 패널에서 조사되고 있는 경제활동 자료는 현재 존재하는 기타 자료에 비하여 매우 상세한 수준까지 조사되고 있다는 점과, 가구 및 개인별 근로소득에 대한 자료를 포함하고 있다는 점에서 매우 유익한 자료라고 할 수 있다.

현재 경제활동 상황에 대한 자료로서 가장 널리 사용되는 자료는 통계청에서 매월 조사·발표하는 「경제활동 인구조사」이며, 근로소득, 특히 임금에 대한 자료로서 가장 널리 사용되는 자료는 노동부에서 매년 조사·발표하는 「임금구조 기본통계조사」라고 할 수 있다. 이와 같은 고용 자료와 임금자료를 사용함에 있어서 그간 문제가 되어 왔던 점은 두 자료가 분리되어 있음으로 인해 개인의 경제활동 상황과 그 대가로 지급되는 임금간에 직접적이고 정확한 연계가 어렵다는 점이었다. 일례로 개인의 노동공급 결정과정을 본인에게 제시되는 임금 및 가구내 소득의 함수로서 추정하는 가장 기초적인 노동공급 함수의 추정도 불가능하였고, 이러한 현실은 학계 및 노동시장에서 공급 측면의 역할이 거의 무시되는 결과를 초래하였다고 하여도 과언이 아니다.

한국 노동연구원 패널자료는 이러한 자료의 결함을 보완함과 동시에, 개별 근로자의 경제활동 상황을 시계열적으로 추적할 수 있는 자료라는 점에서 매우 중요한 의미를 갖는다고 할 수 있다. 이러한 자료의 출현은 향후 학문적인 발전 뿐 아니라, 정책적 시사점을 도출하기 위한 연구에도 높은 유용성을 지닐 것이라고 판단된다. 그러나 한국 노동연구원 패널이 가지고 있는 장점이 최대한 활용되기 위해서는, 이 자료와 현재 존재하는 자료와 일정 수준의 일관성이 유지되어야 할 것이다. 즉, 한국 노동연구원 패널자료의 모집단 대표성이 유지될 필요가 있다는 점이다. 한 자료의 대표성이 유지되지 못할 경우, 그 자료로부터 도출된 결과는 노동시장의 실제 상황을 왜곡되게 표현할 가능성이 높으며, 그러한 결과에 바탕을 둔 노동시장 정책은 상황을 개악할 위험성이 높아지기 때문이다.

본 절에서는 이러한 관점에서 출발하여, 현존하는 자료와 한국 노동연구원 패널의 자료를 비교하여, 각 자료간에 일관성이 유지되고 있는지를 검증하고자 한다. 이러한 분석을 위하여 우선 경제활동(고용) 분석을 위한 자료로서 「경제활동 인구조사」를 사용하며, 임금에 대한 자료로는 「임금구조 기본통계조사」 및 「영세규모 사업체 근로실태조사」를 사용한다. 1998년도를 기준으로 각 자료를 분석한 결과를 비교하며, 이 결과를 바탕으로 자료간 중요한 차이점과 유사점을 밝히고자 한다. 이와 같은 비교에 사용된 자료의 특성은 <표 1>에 간략히 설명되어 있다.

<표 1> 비교에 사용된 자료의 특성

	경제활동인구조사	임금구조기본통계조사	영세규모사업체 근로실태조사
조사주체	통계청	노동부	
조사주기(시점)	매월	매년(3월)	매년(9월)
조사유형	가구조사	사업체 조사	
조사대상	15세 이상 인구	15세 이상 임금근로자	
주요조사항목	인구유형 경제활동유형 근로특성 주당근로시간 산업/직종 종사상 지위 미취업자 실업기간 전직(산업/직종) 실직사유	근로자 유형 근로특성 근로일수 주당근로시간(초과근로 포함) 산업/직종 근속기간 경력기간 임금자료 정액/초과/보너스	

2. 가구 조사 내용의 비교

가. 가구조사에 나타난 인구 특성의 비교

본 절에서는 「경제활동 인구조사」를 이용하여 가구 변수 및 가구주 변수에 대한 비교를 하고자 한다. 주요 비교 변수는 가구원수, 가구내 근로소득과 가구주의 성별, 연령 및 학력 등이다. 이러한 비교의 결과는 가구의 추출이 대표성을 갖는지에 대한 가장 기초적인 근거라고 할 수 있다. 한국 노동연구원 패널과 「경제활동 인구조사」의 가구 및 가구주 변수를 비교한 결과는 <표 2>에 수록되어 있다. 우선 한국노동연구원 패널에서는 5000가구를 대상으로 하고 있는데 반해 「경제활동 인구조사」에서는 월평균 28000여 가구를 조사하고 있다. 따라서 대표성과 표본의 크기에 어느 정도의 상관관계가 있다면, 「경제활동 인구조사」는 효과적인 비교 대상이라고 할 수 있을 것이다.

<표 2>의 결과에 의하면 가구당 15세 이상 가구원 수는 평균적으로 한국 노동연구원 패널에서 다소 높게 나타나고 있다.¹⁾ 한편 가구내 근로소득은 130만원으로 추정되며, 근로소득이 0이 아닌 가구만을 대상으로 할 경우에는 153만원으로 추정되고 있다. 이러한 가구내 근로소득은 「경제활동 인구조사」와 직접적인 비교가 불가능하기 때문에 여기서는 비교를 하지 않았다. 다만 개인별 근로소득은 노동부의 임금자료와 비교가 가능하므로 후에 이에 대하여 자세히 논하기로 한다.

1) 한국 노동연구원 패널은 군인으로 복무하고 있는 경우도 가구원으로 포함시키고 있다. 그러나 이러한 개인의 수는 총 76명에 불과하여, 이를 제외하여도 평균 가구원 수는 2.73명으로 추정되고 있다.

<표 2> 가구수, 가구원수, 가구 근로소득 및 가구주의 인적 특성비교

		한국 노동연구원 패널	경제활동인구조사
가구변수	가구수	5000 가구	28,261 가구(월평균)
	가구원수1)	2.75(3.50) 명	2.47 명
	근로소득2)	130.0(152.9)	-
	근로소득3)	127.9(169.0)	-
	임금소득	88.7(148.7)	-
	비임금소득	39.3(139.0)	-
가구주 변수	남성 비중	85.9%	81.5%
	평균 연령	46.37세	45.3세
	학력분포		
	고졸 미만	38.2%	35.4%
	고졸	39.3%	40.9%
	초대졸	4.8%	4.3%
대졸 이상	17.6%	19.4%	

- 주 : 1) 15세 이상 가구원 수의 평균이며, 괄호 안의 숫자는 15세 미만 가구원을 포함하였을 경우의 평균.
 2) 가구별 자료에서 도출된 가구 총 근로소득(임금+비임금 근로소득)의 평균이며, 괄호안 숫자는 근로소득이 양수인 가구의 평균 근로소득
 3) 개인별 자료에서 도출된 가구 총 근로소득(임금+비임금 근로소득)의 평균이며, 괄호안의 숫자는 근로소득이 양수인 가구의 평균 근로소득
 4) 개인별 자료에서 도출된 가구 임금소득의 평균이며, 괄호안의 숫자는 임금소득이 양수인 가구의 평균 임금소득.
 5) 개인별 자료에서 도출된 가구 비임금 근로소득의 평균이며, 괄호안의 숫자는 비임금 근로소득이 양수인 가구의 평균 비임금 근로소득.

자료 : 한국노동연구원 패널(1998년도), 경제활동인구조사(1998년도).

한국 노동연구원 패널의 소득자료에 있어서 한가지 특이한 점은 가구자료에서 가구내 근로소득을 조사하는 한편, 개인 자료에서 개인별 근로소득을 다시 또 조사한다는 점이다. 이론적으로는 한 가구내 구성원의 근로소득을 모두 합하면 가구내 근로소득과 일치하여야 하므로 이에 대한 검증은 시도하였다. 구성원의 임금 및 비임금 근로소득을 기준으로 가구내 근로소득을 추정할 결과 그 평균 값은 128만원으로, 가구내 근로소득의 평균 값과 약간의 차이를 보이고 있다. 그러나 근로소득이 있는 구성원을 포함한 가구만을 대상으로 평균 가구 근로소득을 추정할 경우에는 169만원으로 추정되어 그 차이가 다소 확대되고 있다. 이러한 차이의 원인은 현재로서 불명확하지만, 실제 가구자료에서 근로소득이 있다고 답한 가구와, 개인별 자료를 토대로 근로소득이 있는 가구와 다소의 차이가 나기 때문인 것으로 판단된다. 이에 대해서는 후에 <표 4>에서 다시 논의하기로 한다.

한편 가구주의 특성을 비교하면, 한국노동연구원 패널의 경우 가구주가 남성인 경우가 다소 많은 것으로 나타나고 있고, 연령에서도 다소 높은 것으로 추정되고 있다. 학력별 분포는 대체로 유사한 편이나, 최저 학력인 고졸 미만의 비중이 다소 높고, 최고학력인 대졸 이상의 비중이 다소 낮게 추정되고 있다. 이는 연령이 높게 추정되는 것과 일정 부분 관련이 있어 보인다. 왜냐하면, 고연령일수록 학력수준이 낮기 때문이다.

나. 가구주의 경제활동 비교

가구주의 경제활동을 비교함에 있어 한국노동연구원 패널과 「경제활동 인구조사」 간에 다소의 정의상 차이가 있음을 밝힐 필요가 있다. 취업의 경우 양 자료가 모두 18시간 미만 무급가족 종사자를 미취업자로 분류하는 점에서 공통점을 갖는다. 반면 실업의 경우 「경제활동 인구조사」에서는 지난 1주간 구직한 미취업자 가운데 즉시 취업이 가능한 경우를 실업으로 정의하고 있으나, 한국노동연구원 패널에서는 지난 4주간 구직 경험이 있는 경우로 확대하여 정의하고 있다. 그러나 한국노동연구원 패널에서는 지난 1주간 구직 한 경우와 2~4주간 구직 한 경우를 분리하여 다루고 있기 때문에 「경제활동 인구조사」와 동일한 정의하에 실업자를 추출할 수 있다. 본 절에서는 이와 같이 실업자를 지난 1주간 구직 한 경우에 한정하여 비교분석을 실행한다.

가구주만을 대상으로 한 경우 경제활동 유형별 분포는 매우 큰 차이를 보이고 있다. 우선 취업자의 비중이 한국노동연구원 패널에서 무려 7.9% 포인트 낮게 추정되고 있으며 이는 편차율이 12.2%에 달하고 있다는 의미이다. 반면 실업자 비중은 3.0% 포인트 높게 추정되고 있어, 편차율이 65.2%에 이르고 있고, 비경제활동 인구의 비중도 4.8% 포인트 높게 나타나 34.0%의 편차율을 보이고 있다. 즉 한국노동연구원 패널에서는 가구주가 취업하고 있지 못할 가능성이 높은 것으로 추정되고, 미취업자 가운데 비경제활동보다는 실업자의 비중이 높은 것으로 나타나고 있다.

이러한 차이의 원인이 가구주의 정의에서 발생될 수도 있다. 「경제활동 인구조사」에서의 가구주는 경제적인 의미에서 정의된 가구주로서, 여성 배우자의 소득이 가구 소득에서 차지하는 비중이 높을 경우 여성 배우자를 가구주로 기록하는 경우가 빈번하다. 이러한 정의상의 차이는 「경제활동 인구조사」에서의 가구주 성별 분포가 여성을 더 많이 포함하고 있음에 일부 반영되고 있다고 할 수 있다. 그러나 이러한 정의상의 차이가 경제활동 유형별 분포에 이와 같이 큰 차이를 초래하였는지는 의문시 된다. 특히 후에 <표 6>과 <표 7>에서 볼 수 있듯이 이러한 차이는 비단 가구주에만 국한되는 것이 아니기 때문에, 자료수집 과정에 근본적인 차이가 존재하는 것으로 판단된다.

<표 3> 가구주의 경제활동상태 비교

(단위 : %)

	취업 또는 일시휴직	실업		비경제활동
		1주간 구직	4주간 구직	
한국노동연구원 패널	73.4	7.6	1.2	17.7
경제활동 인구조사	81.3	4.6	-	14.1

3. 개인자료의 비교

가. 가구자료와 개인자료의 비교

앞서 논의되었던 가구내 근로소득이 가구자료에 의거하였을 경우와 개인자료에 의거하였을 경우의 차이는 일부 응답오차(response error)에 의한 결과일 가능성이 있다. 이를 위해 <표 4>에서는 가구자료를 토대로 근로소득이 존재하는 가구와, 개인자료를 토대로 근로소득의 유무가 파악된 가구의 빈도를 조사하고 있다. <표 4>의 결과에 의하면, 가구조사에서 근로소득이 있다고 응답한 총 3997 가구 가운데 1.3%인 52개 가구에서 근로소득이 있는 구성원이 파악되지 못하고 있다. 한편, 개인자료에서 근로소득이 있는 구성원이 존재하는 것으로 파악된 4252 가구 가운데 7.2%인 307개 가구는 가구 근로소득이 없다고 응답한 가구로 판명되었다.

<표 4> 가구별 근로소득의 유무 및 평균 : 가구자료와 개인자료의 비교

(단위 : 만원, 개)

개인자료 \ 가구자료	근로소득 있음			근로소득 없음		
	개인자료1)	가구자료2)	가구수	개인자료1)	가구자료2)	가구수
근로소득 있음	160.6	156.0	3945	121.7	0	307
근로소득 없음	0	112.7	52	0	0	696

주 : 1) 개인별 자료에서 도출된 가구별 근로소득의 평균.

2) 가구별 자료에서 도출된 가구별 근로소득의 평균.

자료 : 한국 노동연구원 패널, 1998년도.

이러한 응답오차는 비단 빈도에만 존재하는 것이 아니라 소득액수에도 존재하는 것으로 나타났다. 가구자료와 개인별 자료에서 모두 근로소득이 존재하는 것으로 파악된 3945 가구에서도 평균 근로소득액수는 약 4만6천원 가량 차이가 나고 있어 2.9%의 오차율을 보이고 있다. 이 오차율은 그리 크지는 않지만, 자료의 신빙성이라는 측면에서 빈도 및 액수의 오차를 줄이는 노력이 필요한 것으로 판단된다.

나. 인구 구성의 비교

자료의 대표성이라는 측면에서 가장 중요하게 다루어져야 할 부분 가운데 하나가 바로 인구 구성이다. <표 5>에서는 성별, 연령별, 학력별 인구 구성을 한국노동연구원 패널과 「경제활동 인구조사」의 경우를 비교하고 있다. 그 결과에 의하면 양 자료에서 인구구성은 대체로 일치하는 것으로 나타나고 있으나 다소의 차이는 보이고 있다. 우선 한국노동연구원 패널의 경우 남성의 비중이 다소 높게 나타나고 있고, 연령은 청년층의 비중이 대체로 높아 평균 연령도 낮게 나타나고 있다. 한가지 주목할 점은 <표 5>의 결과가 군 복무자를 제외한 결과임에도 불구하고 20~24세 남성의 경우 그 비중이 상대적으로 매우 높

게 나타나고 있다는 점이다. 학력별 분포는 대체로 유사하나 남성의 경우 초대졸 이하의 비중이 다소 높고 대졸이상의 고학력 비중이 다소 낮게 나타나는 점은 가구주의 경우와 유사한 양상을 보이고 있다.

이와 같이 인구 구성에서는 큰 차이를 보이고 있지 않다는 결과는 앞서 본 (가구주의) 경제활동 유형별 분포와 비교하여 중요한 시사점을 갖는다. 즉, 경제활동 유형별 분포에서 보인 큰 차이가 비단 가중치의 문제는 아닐 수 있다는 가능성이 있다. 일반적으로 개인별 자료의 가중치는 인구유형에 의해 결정되는데, 현재 한국노동연구원 패널에는 가중치가 존재하지 않는다. 그러나 가중치를 사용하지 않음에도 불구하고 인구구성이 유사하게 나타난다는 점은 가중치의 필요성이 오히려 높지 않다는 점을 의미할 수도 있다. 이 경우 경제활동 유형별 분포의 차이가 크게 존재한다면, 이는 응답자의 응답 오차(response error) 또는 조사상의 오차(measurement error)일 가능성이 높다. 이러한 오차의 해소는 매우 중요한 문제라고 할 수 있는데, 이 중요성은 다음 개인별 경제활동의 비교에서도 심각하게 대두되고 있다.

<표 5> 인구 구성의 비교

인구 구성지표	한국노동연구원 패널1)		경제활동 인구조사	
	남성	여성	남성	여성
성별 비중	49.2	50.8	48.6	51.4
평균 연령	38.2세	39.5세	39.7세	40.9세
분포				
15~19세	12.6	11.4	10.8	9.5
20~24세	10.2	10.8	6.7	9.6
25~29세	11.4	10.9	12.0	11.4
30~34세	10.8	10.3	11.7	10.7
35~39세	12.0	11.8	13.2	11.6
40~44세	10.9	10.1	11.3	10.5
45~49세	8.2	8.4	8.3	7.4
50~54세	7.2	5.8	6.9	6.4
55~59세	5.8	6.0	6.1	6.0
60세 이상	11.0	14.5	13.0	16.9
학력분포				
고졸미만	33.9	46.8	33.7	47.9
고졸	44.6	39.8	43.7	38.6
초대졸	5.5	5.1	4.7	4.7
대졸이상	16.0	8.4	17.9	8.9

주 : 1) 군 북무자를 제외한 15세 이상 인구를 대상.

자료 : 한국노동연구원 패널(1998년도), 경제활동인구조사(1998년도).

다. 개인별 경제활동의 비교

자료내 모든 구성원을 대상으로 경제활동 유형별 분포를 비교한 결과는 <표 6>에 수록되어 있다. 그 결과에 의하면 앞서 <표 3>에서 본 가구주의 경우와 동일하게 취업비중이 매우 낮고 실업 및 비경제활

동의 비중이 높은 것으로 나타난다. 취업 비중의 경우 무려 10.3% 포인트가 낮게 추정되어 편차율이 18.2%에 이르고 있고, 실업 비중은 3.1% 포인트 높게 추정되어 73.8%의 편차율을 보이고 있다. 결과적으로 실업률은 「경제활동 인구조사」에서는 6.8%로 추정되나, 한국노동연구원 패널에서는 무려 13.5%로 두 배 수준을 보이고 있다.

<표 6> 경제활동 구성비교

	(단위 : %)		
	취업(일시휴직)	실업2)	비경제활동
한국노동연구원 패널1)	46.8	7.3	45.9
경제활동 인구조사	56.5	4.2	39.3

주 : 1) 지난 1주간 구직경험의 유무를 기준으로 구분.

2) 군 복무자를 제외한 15세 이상 인구를 대상.

자료 : 한국노동연구원 패널(1998년도), 경제활동인구조사(1998년도).

<표 7>에 의하면 이러한 양상이 모든 인구 유형에서 동일하게 나타나고 있음을 알 수 있다. 그러나 유형별로 편차의 크기와 편차율이 다소 다르게 나타나고 있는데, 우선 성별로 구분할 때, 취업비중은 남·녀 모든 9% 포인트 정도 낮게 추정되고 있으나 실업비중은 남성의 경우 2.6% 포인트, 여성의 경우 3.8% 포인트 높게 추정되어, 결과적으로 실업률은 각각 4.9% 포인트, 9.7% 포인트 높게 추정되고 있다. 즉, 실업률의 상향편차는 남성보다 여성에게서 보다 심각하게 발생하고 있는 것으로 판단된다.

연령별로는 24세 이하 저연령과 50세 이상 고연령층에서 취업비중의 하향편차가 크게 나타나고 실업 비중의 상향편차도 저연령층과 고연령층에서 크게 나타나고 있다. 결과적으로 실업률의 경우 모든 연령층에서 상향편차가 존재하지만, 성별과 연령별로 다소의 차이를 보인다. <도표 1>에서 볼 수 있듯이 남성의 경우 저연령과 고연령층에서 편차율이 매우 높고 여성의 경우에는 편차율이 연령과 함께 단조 증가하여 50세 이상에서는 300%를 상회하고 있다. 한편 학력별로는 저학력의 취업비중 하향편차 및 실업 비중의 상향편차가 크고, 결과적으로 실업률 편차율은 저학력에서 높게 나타나고 있다.

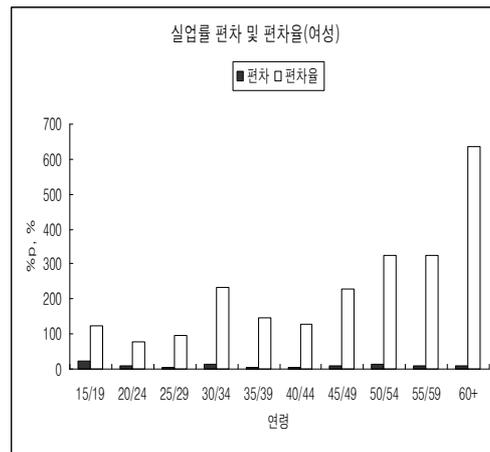
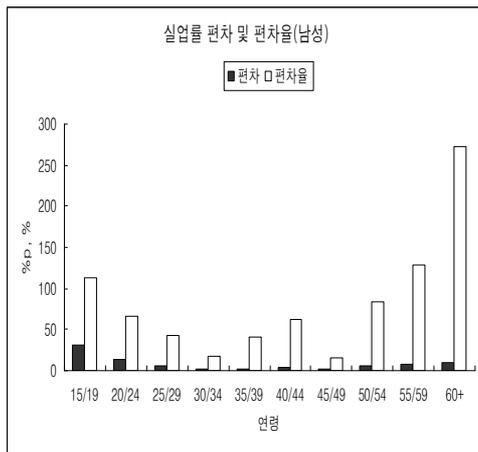
<표 7> 인구 유형별 경제활동 상태

(단위 : %)

인구 구성지표	한국노동연구원 패널1)				경제활동 인구조사			
	남성		여성		남성		여성	
	취업	실업	취업	실업	취업	실업	취업	실업
전체	58.5	8.4	35.8	6.4	69.4	5.8	44.3	2.6
연령별								
15~19세	3.6	5.2	6.6	5.2	5.5	2.1	7.4	1.8
20~24세	20.3	9.8	41.8	11.8	40.1	9.7	51.6	7.3
25~29세	63.5	13.3	40.0	6.1	74.0	10.2	49.9	3.6
30~34세	85.5	6.9	36.1	7.7	90.2	6.1	44.6	2.5
35~39세	86.2	8.1	48.6	6.1	90.6	5.9	54.5	2.6
40~44세	84.1	9.4	54.2	6.6	89.3	5.9	60.1	3.0
45~49세	84.3	7.0	50.5	7.7	88.0	6.2	59.6	2.5
50~54세	76.1	11.3	37.9	7.7	85.9	6.5	52.9	2.2
55~59세	64.6	10.3	40.1	4.8	78.2	5.0	50.5	1.3
60세 이상	36.0	5.2	16.6	1.9	51.3	1.8	28.2	0.4
학력별								
고졸미만	44.0	8.6	32.8	5.2	53.7	4.6	40.6	1.7
고졸	59.8	9.4	34.6	7.7	73.0	7.0	44.9	3.5
초대졸	79.3	8.3	49.0	9.7	85.4	7.9	59.7	5.3
대졸이상	78.7	5.9	50.3	4.4	86.1	4.4	53.8	2.7

주 : 1) 군 복무자를 제외한 인구를 대상.

자료 : 한국노동연구원 패널(1998년도), 경제활동인구조사(1998년도).



<도표 1> 성별, 연령별 실업률 편차 및 편차율

4. 경제활동 유형별 비교

가. 취업자의 취업형태 비교

본 절에서는 취업자만을 대상으로 하여 취업의 형태 및 내용을 비교하기로 한다. <표 8>의 결과에 의하면 취업자만을 대상으로 할 경우 임금·비임금 근로자의 구성은 대체로 일치하는 것으로 나타난다. 전체적으로 취업자 가운데 임금근로자가 차지하는 비중은 한국 노동연구원 패널의 경우 37.6%, 「경제활동 인구조사」의 경우 38.9%로 상당히 유사한 수준이다. 한편 비임금 근로자 가운데 고용주, 자영업 및 무급가족 종사자(주당 18시간 이상 근로)의 비중도 거의 일치하고 있으며, 이러한 양상은 남성 및 여성 취업자에게서 공통적으로 관측되고 있다.

다만 임금 근로자의 구성은 양 자료간에 큰 차이를 보이고 있다. 한국 노동연구원 패널에서는 일반적으로 상용근로자의 비중이 매우 높게 추정되는 반면, 임시직의 비중은 매우 낮게 추정되고 있다. 일용직의 경우는 그 비중이 한국 노동연구원 패널에서 다소 낮게 추정되고 있다. 이러한 결과는 일단 응답자의 응답오차(response error)를 반영하고 있는 것으로 보인다. 즉, 응답자의 경우 상용·임시직의 구분, 또는 임시직·일용직의 구분이 명확하게 이루어지지 못하고 있을 가능성이 존재한다는 의미이다.

한편 기업규모별로는 한국 노동연구원 패널에서 1~4인 규모의 영세기업 비중이 두 배 가량 높게 추정되고 있다는 점이 특이하다. 기업규모에 대한 인식이 비임금 근로자보다 임금근로자에게서 오류가 발생할 확률이 작다는 전제 하에 임금근로자만을 대상으로 규모별 분포를 추정하여도, 그 편이는 다소 작기는 하지만 아직도 상당한 수준에서 존재함을 알 수 있다. 이와 같이 영세규모 사업체의 근로자 비중이 높게 추정되는 경향은 남·녀 모두에게서 공통적으로 나타나고 있어, 이 역시 응답오차 또는 조사상의 오차일 가능성이 높은 것으로 판단된다.

<표 8> 취업자의 취업형태 및 기업규모분포

(단위 : %)

	한국노동연구원 패널			경제활동인구조사		
	전체	남성	여성	전체	남성	여성
고용주	8.1	10.7	3.9	7.2	10.0	3.1
자영업주	21.3	23.5	17.7	21.6	25.2	16.4
무급가족 종사자	8.2	1.6	18.8	10.1	1.7	22.3
상용 근로자	50.4	53.7	45.0	32.3	40.8	19.8
임시직 근로자	6.1	4.7	8.2	20.1	14.9	27.9
일용직 근로자	6.0	5.7	6.4	8.7	7.5	10.6
기업규모1)						
1~4인	24.4(15.7)	22.4(11.8)	28.0(22.2)	48.4(20.4)	42.0(14.0)	58.0(30.6)
5~9인	12.9(12.7)	12.1(11.2)	14.3(15.3)	11.6(15.7)	12.1(15.7)	10.9(16.5)
10인 이상	62.8(71.6)	65.5(77.0)	57.7(62.5)	39.9(63.9)	45.9(20.8)	31.1(52.8)

주 : 1) 전체 취업자의 규모별 분포이며, 괄호 안의 숫자는 임금근로자의 규모별 분포.
 자료 : 한국노동연구원 패널(1998년도), 경제활동인구조사(1998년도).

나. 임금 근로자의 근로 특성

임금 근로자만을 대상으로 근로의 내용 또는 특성(근로일수, 시간 및 근속년수)을 비교한 결과는 <표 9>에 수록되어 있다. 「경제활동 인구조사」에서는 지난 주의 근로시간만이 보고되기 때문에, 보다 자세한 분석을 위하여 본 절에서는 노동부의 「임금구조 기본통계조사」 및 「영세규모 사업체 근로실태조사」를 사용하였다.

<표 9>의 결과에 의하면 전체적으로 월평균 근로일수는 양 자료간에 매우 유사하지만, 주당 근로시간은 한국노동연구원 패널에서 다소 짧게 추정되고 있다. 반면 참고로 기입된 「경제활동 인구조사」의 평균 근로시간에 비해서는 오히려 긴 편으로 추정되고 있다. 한편 평균 근속년수는 양 자료에서 매우 유사하게 약 60개월 수준에서 추정되고 있다는 점이 특이하다.

<표 9> 임금 근로자의 성별·기업규모별 근로 특성

(단위 : %)

	한국노동연구원 패널			임금구조 기본통계조사1)		
	월근로일수	주당근로시간	근속개월	월근로일수	주당근로시간2)	근속개월
전체	24.6	50.6	60.2	24.4	53.1(48.9)	61.1
기업규모별						
1~4인	24.9	55.2	33.2	25.9	54.1(48.4)	31.3
5~9인	24.5	51.9	43.4	25.5	51.2(49.2)	44.7
10인 이상	24.6	49.4	68.8	23.7	53.0(49.0)	72.7
남자	24.6	52.1	71.7	24.1	54.0(50.5)	71.2
1~4인	24.9	57.5	40.4	25.8	53.8(50.2)	39.1
5~9인	24.1	54.0	50.8	25.5	51.9(50.3)	49.3
10인 이상	24.6	51.0	78.9	23.6	54.3(50.6)	80.8
여자	24.6	48.0	40.4	24.9	51.3(46.3)	40.3
1~4인	24.8	53.0	26.6	26.0	54.2(47.0)	23.1
5~9인	24.9	49.2	33.8	25.6	50.0(47.6)	35.9
10인 이상	24.5	45.9	46.9	24.1	49.8(45.5)	51.1

주 : 1) 임금구조기본통계조사 및 영세규모사업체 근로실태조사에서 추출된 자료,
2) 괄호 안의 숫자는 경제활동인구조사에서 추출된 자료.

자료 : 한국노동연구원 패널, 임금구조기본통계조사, 영세규모사업체 근로실태조사, 경제활동인구조사, 각 1998년도.

한편 임금근로자를 기업 규모별로 구분할 경우 다소 차이를 보인다. 우선 노동부의 자료에 의하면 기업규모가 증가함에 따라 월 근로일수는 단조 감소하는 반면, 주당 근로시간은 특정한 패턴을 보이고 있지 않다. 이와 반대로 한국노동연구원 패널에서는 기업규모가 증가함에 따라 월 근로일수가 상대적으로 일정하고, 주당 근로시간이 감소하는 양상을 보인다. 결과적으로 양 자료에서 총 근로시간(근로일수*근

로시간/7)은 공통적으로 기업규모에 따라 감소하고 있다. 한편 근속년수는 양 자료간에 매우 유사한 수준을 보이고 있다.

다. 임금근로자의 임금구조

임금 근로자만을 대상으로 임금구조를 비교한 결과는 <표 10>에 수록되어 있다.²⁾ 이 결과에 의하면 한국노동연구원 패널에서 전체적으로 임금이 다소 높게 추정되고 이는 남성의 임금이 높게 추정된 결과를 반영하고 있다. 결과적으로 성별 임금격차는 한국노동연구원 패널에서 다소 높게 추정되고 있다.

<표 10> 민간 부문 임금 근로자 유형별 임금

(단위 : 만원)

	한국노동연구원 패널		임금구조 기본통계조사1)	
	남성	여성	남성	여성
전체	130.8 (65.7)2)	77.4 (48.5)	122.3 (56.3)	77.5 (33.9)
연령별				
15~19세	57.4 (17.9)	58.5 (20.5)	63.2 (29.7)	63.2 (21.7)
20~24세	80.1 (28.6)	76.0 (20.5)	76.0 (24.3)	68.5 (19.6)
25~29세	110.3 (45.1)	87.9 (38.7)	96.7 (28.6)	81.0 (25.6)
30~34세	132.4 (49.1)	89.6 (53.4)	118.4 (38.6)	92.0 (39.1)
35~39세	146.7 (56.7)	82.4 (58.3)	135.2 (49.1)	85.9 (44.8)
40~44세	158.0 (73.2)	73.6 (45.0)	145.4 (58.8)	80.9 (41.6)
45~49세	158.2 (92.4)	68.3 (36.3)	148.6 (70.2)	79.7 (41.9)
50~54세	138.4 (79.9)	80.4 (89.2)	144.5 (77.9)	76.0 (41.9)
55~59세	114.5 (69.3)	54.4 (23.1)	127.2 (80.5)	72.2 (45.8)
60세 이상	95.8 (79.9)	51.8 (19.2)	107.8 (86.6)	72.1 (54.4)
학력별				
고졸미만	98.7 (42.1)	60.3 (36.2)	103.1 (39.1)	69.9 (24.6)
고졸	118.5 (48.0)	78.7 (49.9)	114.1 (46.4)	75.7 (28.6)
초대졸	130.8 (63.8)	81.8 (34.0)	125.1 (52.2)	84.2 (32.0)
대졸이상	177.7 (83.0)	109.3 (58.3)	160.0 (73.3)	123.5 (63.8)
기업규모별				
1~4인	94.2 (40.1)	62.2 (25.2)	103.7 (40.9)	71.7 (24.1)
5~9인	110.3 (45.7)	72.9 (35.7)	115.8 (50.3)	77.0 (30.4)
10인 이상	139.8 (68.8)	84.1 (55.3)	127.2 (58.9)	80.9 (38.5)

주 : 1) 임금구조 기본통계조사 및 영세규모 사업체 근로실태 조사의 자료에서는 연간 보너스를 제외한 임금을 사용.

2) 괄호 안의 숫자는 임금의 표준 편차임.

자료 : 한국노동연구원 패널, 임금구조 기본통계조사, 영세규모 사업체 근로실태 조사, 각 1998년도.

2) 근로소득의 경우 임금소득의 평균은 111.8만원, 비임금 근로자의 소득 평균은 133.1만원으로 추정되어 비임금 근로자의 평균소득이 20%가량 높은 것으로 추정되고 있다.

한편 연령별로는 매우 다른 양상을 보이고 있다. 우선 남성의 경우 20~49세의 임금이 높게 추정되고 나머지 연령층에서는 오히려 임금이 낮게 추정되고 있다. 여성의 경우에는 20~29세의 젊은 여성근로자의 임금이 높게 추정되고 있는 반면, 10대나 30세 이상에서는 임금이 오히려 낮게 추정되고 있다. 이러한 차이에도 불구하고, 연령별 임금구조(age-earnings profile)에 있어서는 양 자료에서 공통적으로 남성의 경우 45~49세에, 여성의 경우에는 30~34세에 임금이 최고점에 이르는 것으로 나타나고 있다.

학력별로는 남성의 경우 고졸미만의 경우에는 임금이 낮게 추정되고 있으나 고졸 이상의 학력에서는 임금이 높게 추정되고 있고, 그 편차 및 편차율은 고학력일수록 높아 대졸 이상의 경우에는 편차가 17.7만원(편차율 : 11.1%)에 이르고 있다. 결과적으로 남성 근로자의 학력별 임금격차는 한국 노동연구원 패널에서 높게 추정되고 있다. 반면 여성의 경우에는 고졸 여성의 임금이 높게 추정되는 반면 다른 학력에서는 오히려 임금이 낮게 추정되고 있어 대졸/고졸 임금 격차가 오히려 낮게 추정되고 있다.

기업의 규모에 의해서는 임금이 단조 증가하는 추세를 보이고 있어 1~4인 기업의 평균은 76.7만원, 5~9인 평균은 93.6만원, 그리고 10인 이상 평균은 121.6만원에 이르고 있다. 이러한 유형은 노동부의 임금자료에서도 관측된다. 한편 성별로 구분하여도 임금의 기업규모 효과는 모두 뚜렷하게 나타나고 있는데, 다만 규모별 임금격차는 다소 크게 추정되고 있다. 특히 이러한 경향은 여성에게서 높다. 1~4인 기업과 10인 이상 기업을 비교한 임금격차의 경우 한국 노동연구원 패널에서는 남녀 각각 48.4%와 35.2%로 추정되고 있고, 노동부의 임금자료에서는 22.7%와 12.8%로 추정되고 있다. 즉, 규모별 임금격차는 남성의 경우 약 두 배로 추정되는 반면 여성은 세 배까지 높은 수준으로 추정되고 있다.

라. 임금근로자의 임금분포

마지막으로 임금근로자의 경우 임금소득의 불평등도는 양 자료에서 매우 유사하다. <표 11>의 결과에 의하면 전체적으로 임금분포의 Gini Coefficient는 한국 노동연구원 패널에서 0.298, 노동부의 임금자료에서는 0.296으로 추정되어 거의 차이가 없다. 이러한 결과는 기업 규모별 근로자 분포가 매우 큰 차이를 보였던 점(<표 8> 참조)과 규모별 임금격차가 매우 높게 추정되었던 점(<표 10> 참조)을 감안하면, 이러한 결과는 다소 의외라고 할 수 있다.

<표 11> 민간 부문 임금근로자의 기업 규모별 임금분포(Gini Coefficient)

	전체	기업규모별		
		1~4인	5~9인	10인 이상
한국노동연구원 패널	0.298	0.253	0.251	0.291
임금구조 기본통계조사	0.296	0.239	0.266	0.288

자료 : 한국노동연구원 패널, 임금구조 기본통계조사, 영세규모 사업체 근로실태 조사, 각 1998년도.

이와 같이 기업규모별 근로자 분포와 규모별 임금격차는 매우 상이하면서도 전반적인 임금분포가 매우 흡사한 원인은 기업 규모별 임금분포의 차이에서 찾을 수 있다. <표 11>에 의하면 소규모 기업일수

록 임금분포가 평등한 양상은 양 자료에서 공통적으로 나타나고 있으나, 대체로 각 규모별로 임금 분포의 Gini Coefficient는 작게 추정되는 경우가 더 많다. 즉 기업규모간 평균임금의 격차는 크지만, 규모 내에서 분포는 상대적으로 집중되어 있는 점이 서로 상쇄되어 전반적인 임금분포의 불평등도는 유사하게 추정되고 있는 것이다.

마. 실업자의 구성 비교

이제 실업자의 구성에 대한 비교결과를 알아보자. <표 12>는 한국노동연구원 패널과 「경제활동 인구조사」에서 추출된 실업자의 구성을 비교하고 있다. 이 결과에 의하면 실업자 가운데 여성의 비중이 높게(11.2% 포인트) 평가되고 있어, 앞서 <표 7>에서 보았던 결과와 일관성을 보여주고 있다. 다만 실업자 범위를 4주 이내 구직 경험자로 넓힐 경우 남성비중이 다소 증가하고 있어, 2~4주 이내 구직 경험자 가운데는 남성의 비중이 높음을 의미하고 있다. 그러나 이 결과는 1주간 구직여부만을 공표하는 「경제활동 인구조사」와 직접적인 비교가 불가능하다.

한편 연령별로는 전반적으로 여성의 비중이 크기 때문에 전체 실업자에서 차지하는 비중이란 측면에서는 모든 연령층 여성의 비중이 높게 추정되고 있다. 각 성별 내에서 연령별로는 남성의 경우, 25세 미만, 50세 이상의 청·노년층의 비중이 크게 추정되고 있어 <도표 1>에서 보았던 실업률 편차 양상과 동일한 양상을 보이고 있다. 여성의 경우에도, 20세 미만, 45세 이상의 비중이 높게 추정되어 <도표 1>의 실업률 편차 양상과 대체로 동일한 양상을 보인다. 학력별로는 남녀 모두에게서 공통적으로 저학력의 비중이 높게 추정되고 대졸 이상의 고학력에서는 비중이 낮게 추정되고 있다. 이 결과 역시 <표 7>의 결과와 동일한 양상을 보이고 있다.

바. 실업기간(구직기간)의 비교

<표 13>은 한국노동연구원 패널 및 「경제활동 인구조사」에서 구직기간을 평균하여 비교하고 있다. 이 결과에 의하면 실업기간은 한국 노동연구원 패널에서 매우 길게 추정되고 있으며 이러한 양상은 남성보다는 여성에게서 더욱 심하게 나타나고 있다. 「경제활동 인구조사」에서는 평균 구직기간이 3.13개월에 불과한데 반하여 한국노동연구원 패널에서는 8.14개월로 추정되고 있어 거의 2.6배에 이르고 있다. 한편 남자의 경우 구직기간이 4.54개월(135%) 길게 추정, 여성의 경우에는 6.21개월(329%) 길게 추정되고 있다.

연령별로는 남녀 모두에게서 일반적으로 중·장년층의 상향편차가 매우 심하게 나타나고 있고, 학력별로는 고학력·장기실업 양상이 양 자료의 공통점으로 유지되고는 있으나 남성의 경우 고학력의 상향편차가 심하게 나타나고 있다. 여성의 경우에는 실업기간의 상향편차가 학력과 관계없이 대체로 일정한 편이다.

이와 같이 실업기간이 길게 추정되는 것은 결과적으로 실업비중 차이(<표 6> 참조)의 상당부분을 실

업기간이 설명할 수 있는 가능성을 제시하고 있다. 만일 실업기간이 경제활동 인구조사와 같이 3.13개월 이라고 가정한다면, <표 6>에서 실업비중은 7.3%가 아니라 $2.7\%(=7.3 \times 3.13 / 8.14)$ 가 되었을 것이고, 따라서 실업률은 13.5%가 아니라 5.5%이었을 것이라는 추정이 가능하다. 즉, 한국 노동연구원 패널에서 실업비중이 높은 것은 기본적으로 응답자들이 구직기간과 미취업기간을 명확하게 구분하지 못하고 있음으로 인해 유발된 결과로 이해할 수도 있을 것이다.

<표 12> 실업자 유형별 구성

(단위 : %)

	한국노동연구원 패널1)		경제활동 인구조사	
	남성	여성	남성	여성
전체	56.3 (53.9)	43.8 (43.8)	67.4	32.6
연령별				
15~19세	7.8 (9.0)	9.3 (9.5)	4.0	6.6
20~24세	11.8 (13.3)	20.0 (21.1)	11.3	26.7
25~29세	17.8 (17.0)	10.4 (11.8)	21.2	15.8
30~34세	8.8 (8.7)	12.5 (11.4)	12.4	10.0
35~39세	11.5 (10.2)	11.3 (11.4)	13.5	11.7
40~44세	12.2 (11.2)	10.4 (10.6)	11.5	11.8
45~49세	6.7 (6.2)	10.2 (9.5)	8.8	7.0
50~54세	9.7 (9.6)	7.0 (6.2)	7.8	5.3
55~59세	7.1 (7.1)	4.5 (4.3)	5.4	2.9
60세 이상	6.7 (7.5)	4.3 (4.2)	4.1	2.3
학력별				
고졸미만	33.9 (33.1)	38.1 (36.7)	26.9	30.2
고졸	49.6 (50.7)	48.3 (49.3)	53.1	51.2
초대졸	5.5 (5.2)	7.7 (7.8)	6.4	9.6
대졸이상	11.1 (11.0)	5.9 (6.2)	13.6	9.0

주 : 1) 1주 이내 구직 경험자를 나타내며, 4주 이내 구직경험자를 포함한 수치는 괄호 안에 표기
 자료 : 한국노동연구원 패널, 경제활동 인구조사, 각 1998년도.

5. 맺음말

본 절에서는 한국 노동연구원 패널을 「경제활동 인구조사」, 「임금구조 기본통계조사」 및 「영세 규모 사업체 근로실태조사」에 비교분석하여 대표성을 검증하고자하였다. 일반적으로 한국 노동연구원 패널의 경우 취업비중이 매우 낮게 추정되는 반면 실업비중과 비경제활동 인구의 비중은 매우 높게 추정되는 문제가 심각하게 존재하고 있음이 파악되어, 이에 대한 시정이 시급한 것으로 드러났다. 또한 이러한 양상이 성/연령/학력별로 차이를 보이고 있어 전반적으로 대표성을 유지하기 위해서는 가중치 또는 표본 설계상의 개선이 필요한 것으로 판단된다. 반면 취업자에 국한하여 취업형태 및 근로내용을 비

교한 결과, 상당부분 기존의 자료와 일치된 양상을 보이고 있는 것으로 나타났다. 특히 노동부의 임금자료와는 유사점이 차이점보다 많은 것으로 보인다. 결과적으로 가장 시급한 문제라고 판단되는 것은 경제활동 유형별 분포의 차이점이라고 판단된다.

<표 13> 평균 구직기간의 비교

(단위 : 개월)

	한국노동연구원 패널1)		경제활동 인구조사	
	남성	여성	남성	여성
전체	7.87	8.92	3.33	2.71
연령별				
15~19세	3.52	2.81	2.45	2.12
20~24세	7.93	6.32	3.03	3.03
25~29세	7.00	12.15	3.78	2.92
30~34세	8.89	10.56	3.38	2.64
35~39세	5.70	11.40	3.40	2.49
40~44세	8.12	8.08	3.27	2.42
45~49세	12.16	6.57	3.72	2.81
50~54세	8.55	12.33	3.69	2.84
55~59세	9.63	7.24	3.18	2.04
60세 이상	9.71	18.87	3.16	2.40
학력별				
고졸미만	7.33	9.42	3.04	2.50
고졸	7.10	8.23	3.21	2.61
초대졸	11.37	9.59	3.89	3.34
대졸이상	11.13	10.48	4.07	3.31

주 : 1) 1주 이내 구직 경험자만을 대상으로 평균 구직기간을 추정.
 자료 : 한국노동연구원 패널, 경제활동 인구조사, 각 1998년도.

III. 패널조사에서 표본 이탈자의 특성

패널조사를 위해 선정한 최초의 표본이 대표성을 가진다고 하더라도 이후의 계속된 조사에서 표본의 이탈이 클 경우 확보된 대표성을 상실할 수 있다. 이는 패널자료의 가치를 떨어뜨리는 가장 큰 요인이다. 제3장은 국내에서 구축된 기존의 패널조사 자료를 사용하여 표본 이탈자들의 특성을 분석함으로써 패널조사에서 표본 이탈률(attrition rate)을 최소화하기 위해서는 어떤 노력들이 필요한지에 대한 시사점을 얻고자 한다. 이러한 시사점들은 한국노동패널(KLIPS)이 성공적인 패널조사로 정착되도록 하는 데에 기여할 것이다.

1. 기존의 패널자료

한국노동패널(KLIPS)은 1차 조사가 마무리된 상황이므로 표본 이탈자들의 특성을 분석하기 위해 이 조사 자료를 사용할 수 없다. 따라서 본 연구는 우리나라를 대상으로 실시한 기존의 패널조사 자료를 사용하여 표본 이탈자들의 특성을 분석하기로 한다.

기존의 패널자료는 (1) 대우경제연구소에서 1993년부터 1998년까지 매년 패널조사를 통해 구축하고 있는 「한국가구패널(KHPS)」³⁾과 (2) 통계청에서 실시하는 「경제활동인구조사」를 사용하여 1981년부터 1994년까지 14년에 걸쳐 월별로 패널자료를 구축한⁴⁾ 남재량(1997)의 패널을 시작으로 남재량·이창용(1998), 그리고 남재량·류근관(2000)에 의해 확장되고 있는 패널이 대표적이다.

대우경제연구소에서 실시한 「한국가구패널」의 표본 이탈률은 최초 표본에 비해 18.1%, 28.4%, 35.7%로서 4차년도 조사에서는 최초 표본의 64.3%만이 표본으로 유지되고 있다. 처음 실시된 조사에서 선정한 표본의 무응답을 고려하면 4차년도에 조사된 표본가구는 최초에 임의로 선정한 표본가구의 1/3에 불과하다⁵⁾. 이러한 표본이탈의 문제는 외국의 패널조사에서도 마찬가지로 나타나는 문제이지만, 미국의 PSID나 NLS에 비해 대우 패널조사의 표본 이탈률이 매우 높다. 대우에서 실시한 3년간의 패널조사에서 발생한 표본 이탈률은 PSID의 12년 동안에 걸친 표본 이탈률에 해당한다.

반면 「경제활동인구조사」를 사용하여 구축한 패널자료는, 조사가 매월 실시된다는 점에서 「대우패널」과 다르며 최초 추출되어 조사된 표본의 2차 조사 응답률이 1990년대에는 95%을 훨씬 넘고 있다는 특징이 있다⁶⁾. 따라서 본 연구는 경제활동인구조사를 사용한 패널자료와 대우패널을 모두 사용하여 분석함으로써 「한국노동패널(KLIPS)」의 표본 이탈 최소화를 위한 노력에 대한 보다 풍부한 시사점을 얻고자 한다.

2. 경제활동인구조사를 사용한 패널자료에서 표본 이탈자의 특성

제2절에서는 먼저 경제활동인구조사를 사용하여 표본 이탈자의 특성에 대해 분석한 다음 대우패널을 사용하여 유사한 분석을 하기로 한다. 먼저 경제활동인구조사 자료를 본 연구의 분석에 알맞은 패널자료의 형태로 구축하는 과정에 대해 간략히 설명하고 다음으로 표본 이탈자들의 특성에 대해 분석하기로 하자.

가. 경제활동인구조사를 사용한 패널자료

3) 「한국가구패널」은 「대우패널」로 불리기도 한다. 이하에서 언급되는 「대우패널」은 대우경제연구소에서 실시한 「한국가구패널(KHPS)」을 의미한다.

4) 단, 1981년과 1982년은 분기별이다.

5) 신동균, 『미국 패널데이터의 현황과 시사점 — PSID, NLSY, KHPS, KLIPS를 중심으로 —』, 한국노동패널연구 97-02, 1998. 4.

6) 남재량, 『한국의 실업률 추세변화에 관한 연구』, 1997. 8. p.57, p.88.

한국노동패널(KLIPS)이 연 1회 조사를 실시하는 패널이므로 경제활동인구조사를 사용한 패널도 매년 동일 시점(月)을 기준으로 연도별 패널의 형태로 구축하는 것이 좋다. 경제활동인구조사는 최초 선정된 표본을 평균적으로 5년 동안 유지하고 있는데, 이는 인구주택총조사(센서스)를 모집단으로 하고 있고 센서스가 5년마다 실시되기 때문이다.⁷⁾ 1990년대 들어 경제활동인구조사의 표본은 1993년 2월과 1998년 1월에 개편된 바 있으므로 분석에 사용할 수 있는 자료는 1993년부터 1997년까지 자료와 1998년부터 1999년까지 자료이다. 그러나 후자의 경우는 1차 조사인 1998년 1월과 2차 조사인 1999년 1월 두 연도에 걸친 패널자료만 가능하다는 문제가 있다. 반면 1993년 2월에 개편된 경제활동인구조사 자료를 사용하면 1993년 2월부터 1997년 2월까지 총 5차에 걸친 패널자료를 얻을 수 있다.⁸⁾ 따라서 본 연구는 1993년 2월에 개편된 표본을 사용하여 1997년 2월에 이르기까지 매년 2월 자료를 사용하여 패널자료를 구축하고 이를 분석에 사용하기로 한다.⁹⁾

분석에 사용되는 변수들은 표본 및 표본 이탈자의 성, 연령, 교육수준, 혼인상태, 가구주와의 관계, 경제활동상태, 그리고 종사상의 지위 등이다. 그리고 표본에서 이탈한 사람들에 대한 정보는 더 이상 기록되지 않으므로 이들에 대한 정보는 전년도 조사 자료로부터 구해야 한다. 즉 분석에 사용되는 표본 이탈자들의 특성은 전년도에 실시된 조사에 기록된 정보를 바탕으로 하고 있다.

나. 이탈되는 표본의 특성

최초 선정된 표본에서 이탈되는 표본이 어떠한 특성을 가지는가에 대한 분석은 크게 두 가지로 나누어 진행된다. 첫째는 표본에서 이탈한 사람들 모두를 대상으로 한 분석이다. 이를 통해 표본 이탈로 인한 대표성의 왜곡이 어느 정도이며 어떤 형태로 나타나는가를 살펴볼 수 있다. 둘째는 표본에서 이탈된 가구에 대한 분석이다. 경제활동인구조사는 KLIPS나 대우패널과 마찬가지로 가구조사이므로 가구에 소속된 개별 가구원들에 대한 분석 외에 가구 자체가 표본에서 이탈하는 문제가 중요하다.

1) 표본 이탈과 패널자료의 대표성

【표본의 수와 표본 이탈률】

표본 이탈자들의 구체적인 특성을 보기에 앞서 분석에 사용되고 있는 표본의 수와 표본 이탈자의 수

7) 남재량(1997), pp.52-53.

8) 실제로는 조사가 매월 실시되므로 1년 동안에만 12회의 조사가 이루어진다. 그러나 본고에서 구축되어 사용될 패널은 연도별 패널이므로 편의상 1992년 2월의 조사를 1차 조사, 1994년 2월 조사를 2차 조사, 그리고 1997년의 조사를 5차 조사 등으로 부르기로 한다.

9) 경제활동인구조사를 사용한 패널자료 구축과정에 대한 자세한 설명은 남재량(1997), 남재량·이창용(1998), 또는 남재량·류근관(2000)을 참조하라.

및 표본 이탈률에 대해 살펴보자. <표1>은 이를 정리한 것이다. 표를 통해 알 수 있는 것은 2차 조사에서 이탈률이 가장 높으며 조사가 거듭될수록 표본 이탈률이 하락한다는 사실이다.

1993년 2월에 표본으로 선정되어 1차 조사에서 조사된 사람의 수는 78,465명이다. 이 중 2차 조사인 1994년 2월에도 역시 표본에 남아있는 사람은 53,394명으로 1년 사이에 표본의 32%인 25,071명이 표본에서 이탈하였다. 3차 조사인 1995년 2월에는 1994년 2월에 남아있던 표본의 22.8%인 12,184명이 표본에서 이탈하여 33,050명이 계속 표본으로 확보되고 있다. 22.8%는 여전히 높은 이탈률이지만 2차 조사에 비하면 이탈자수와 이탈률이 크게 하락한 것이다. 4차 조사인 1996년에 이탈자수는 8,160명, 이탈률은 19.8%이고 1997년에는 이탈자수와 이탈률이 각각 5,315명과 16.1%로 감소하고 있다.¹⁰⁾ 즉 조사가 거듭될수록 이탈률이 하락하고 있으며, 특히 3차 조사에서 이탈률의 하락폭이 9.2% 포인트로서 4차 조사의 3.0% 포인트, 5차 조사의 3.7% 포인트 보다 더 크다.

이처럼 조사가 여러 차례 진행되기 전에 표본의 이탈이 매우 크다는 사실은 패널조사에 선정된 표본의 대표성이 크게 훼손될 가능성이 있음을 암시한다. 따라서 표본의 이탈을 줄여 대표성을 유지함으로써 패널조사가 성공적으로 정착되도록 하기 위해서는 패널조사가 여러 번 진행되기 전인 2차 또는 3차 조사에서 표본 이탈자 수를 줄이기 위한 노력이 집중되어야 할 것이다.

<표1> 표본 이탈자 및 이탈률

(단위 : 명, %)

구 분	1993년	1994년	1995년	1996년	1997년
표 본 수	78,465	53,394	41,210	33,050	27,735
이 탈 자 수		25,071	12,184	8,160	5,315
이 탈 률		32.0	22.8	19.8	16.1

자료 : 경제활동인구조사 원자료

【성별 특성】

이제 표본에서 이탈하는 사람들은 주로 어떤 특성을 가지는 사람들인지를 살펴보자. 먼저 표본 이탈자들이 성별로 어떠한 구분되는 특징을 보이는지 보기로 한다. <표2>는 1993년 2월에 선정되어 조사된 표본의 수와 표본 이탈자의 수 및 이탈률을 성별로 보여주고 있다. 표에서 보는 바와 같이 2차 조사에서 남성의 이탈률이 34.3%로서 여성의 이탈률 29.9% 보다 높다. 이후의 거듭된 조사에서 남성과 여성의 이탈률이 모두 지속적으로 하락하고 있으며 이탈률의 성별 격차도 다소 감소하고 있다.¹¹⁾ 그러나 남성이 여성 보다 이탈률이 높은 현상은 여전히 있다. 이처럼 이탈률에 성별 차이가 존재한다는 사실은 최초

10) 이러한 표본 이탈로 최초로 선정되어 조사된 1차 조사 연도의 표본 78,465명 가운데 5차 조사까지 남아있는 표본은 원래 표본의 35.3%인 27,735명이다.

11) 특히 3차 조사에서 이탈률의 성별격차(3.1% 포인트)가 2차 조사(4.4% 포인트)에 비해 상당한 정도로 감소하고 있다.

표본 추출률의 역수를 곱하여 성별 패널자료를 생성시킬 경우 그러한 자료의 신뢰성이 크게 손상됨을 의미한다. 단, ‘표본 및 이탈자수’에서 ‘1993년’은 1993년 2월에 조사된 표본의 수이며 ‘1994년’부터 ‘1995년’까지는 표본에서 이탈한 사람의 수를 나타낸다¹²⁾.

<표2> 이탈자의 성별·조사회수별 특성

(단위 : 명, %)

구 분		1993년	1994년	1995년	1996년	1997년
표본 및 이탈자수	남성	36,459	12,514	5,873	3,858	2,501
	여성	42,006	12,557	6,311	4,302	2,814
이 탈 률	남성		34.3	24.5	21.3	17.6
	여성		29.9	21.4	18.6	14.9

자료 : 경제활동인구조사 원자료

【연령별 특성】

다음으로 표본에서 이탈한 사람들이 연령별로 어떠한 특징을 보이는지 살펴보도록 하자. <표3>과 <그림1>은 그러한 특징을 요약한 것이다. 표와 그림에서 보는 바와 같이 이탈률은 연령에 따라 큰 차이를 보이고 있으며 대체로 (1) 연령이 높을수록, 그리고 (2) 조사가 거듭될수록 이탈률이 하락하는 모습을 보인다.

먼저 연령이 높아질수록 이탈률이 하락하는 특징에 대해 살펴보자. 2차 조사에서 15-24세와 25-34세 연령의 이탈률은 각각 39.8%, 42.8%로 매우 높으며¹³⁾ 연령이 35-44세, 45-54세, 그리고 55-64세로 높아짐에 따라 이탈률은 각각 28.2%, 20.4%, 그리고 18.5%로 낮아진다. 이후의 조사에서는 15-24세의 이탈률이 25-34세의 이탈률 보다 높아 연령이 높을수록 이탈률이 하락하는 특징이 더욱 강하게 나타난다. 다만 나이가 많아 사망확률이 높은 65세 이상 연령집단은 45-54세와 55-64세 연령보다 이탈률이 높다.

다음으로 조사가 계속됨에 따른 이탈률 하락에 대해 살펴보자. 이러한 특징은 모든 연령층에서 공통적으로 나타나는 현상이지만 특히 25-54세의 基幹 노동력(primary labor force)에서 강하게 나타난다. 이 탈률 하락추세가 가장 강한 25-34세를 보면 2차 조사에서 42.8%로 높던 이탈률이 3차 조사에서는 33.0%로, 4차 조사에서는 29.2%로, 그리고 5차 조사에서는 22.2%로 하락하고 있다.

그리고 15-24세와 65세 이상 연령의 이탈률 하락은, 다른 연령 집단에 비해, 4차와 5차 조사에서 그리 크지 않으며 각각 31%와 12% 수준에서 더 이상 하락하지 않고 있다. 15-24세 연령층은 청소년 집단으로서 대개 ‘통학’으로 분류되는 비경제활동상태에 있으며 군 입대나 학교를 마친 후 노동시장에 진입하

12) 이는, 특별한 언급이 없는 한, 이후에 제시되는 표에서도 마찬가지이다.

13) 나이가 적은 연령층이 나이가 많은 연령층 보다 이탈률이 낮게 나타나는 것은 2차 조사의 15-24세와 25-34세에서 유일하다.

는 연령에 해당한다. 이에 따라 동일 가구 내의 다른 가구원들과 지리적으로 떨어져 생활하게 되는 경우가 많이 발생할 것이므로 이들 연령층에 속하는 가구원들의 가구 정착도는 상대적으로 약할 것이다. 이에 따라 이 연령층은 표본 이탈률의 하한선이 다른 연령층 보다 더 높고 따라서 이탈률의 하락폭이 미약한 것으로 생각된다. 65세 이상 연령층의 이탈률이 3차 이후 조사에서 크게 하락하지 않는 것은 이들의 높은 사망확률과 밀접하게 관련되어 있을 것으로 생각된다.

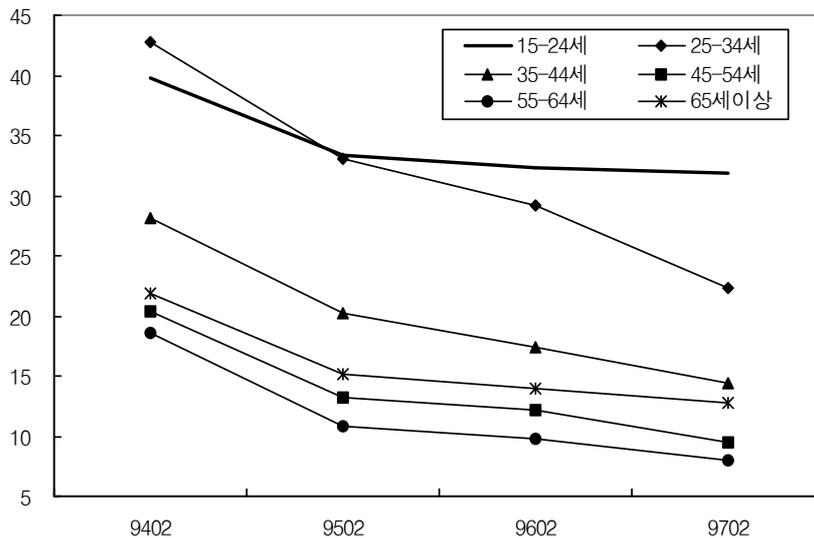
이처럼 표본 이탈률이 연령별로 큰 차이를 보이고 조사가 거듭됨에 따른 이탈률 변화에도 차이가 있으므로 패널조사 결과를 사용한 각종 연령별 통계치들은 상당한 편의(bias)를 가지게 될 것이다.

<표3> 이탈자의 연령별 · 조사회수별 특성

(단위 : 명, %)

구	분	9302	9402	9502	9602	9702
표본 및 이탈자수	15-24세	18,124	7,221	3,642	2,347	1,564
	25-34세	19,642	8,409	3,708	2,198	1,185
	35-44세	15,249	4,296	2,208	1,517	1,038
	45-54세	10,941	2,233	1,152	924	631
	55-64세	7,921	1,468	697	566	416
	65세이상	6,588	1,444	777	608	481
이탈률	15-24세		39.8	33.4	32.3	31.8
	25-34세		42.8	33.0	29.2	22.2
	35-44세		28.2	20.2	17.3	14.4
	45-54세		20.4	13.2	12.2	9.5
	55-64세		18.5	10.8	9.8	8.0
	65세이상		21.9	15.1	13.9	12.8

자료 : 경제활동인구조사 원자료



<그림1> 이탈률의 연령별 · 조사회수별 특성

【교육정도별 특성】

표본에서 이탈한 사람들을 교육수준별로 보면, <그림2>에서 잘 볼 수 있는 바와 같이, 교육수준이 높을수록 이탈률이 높아지는 경향이 강하게 나타난다. <표3>에서 2차 조사의 이탈률은 초등학교 졸업이하의 학력수준에서는 21.4%, 중졸은 27.8%로 높지 않으나 고졸의 경우 36.4%로 높으며 초대·전문대에서는 40.5%, 그리고 대졸이상에서는 41.1%로 매우 높다. 3차 조사부터는 초대·전문대와 대졸이상의 이탈률이 역전되는 현상이 나타나기는 하지만 여전히 학력이 높을수록 이탈률이 높아지는 현상이 관찰된다. 교육정도가 높을수록 이탈률이 높게 나타나는 현상은 학력간 임금격차와 밀접한 관계를 가지는 것으로 판단된다.

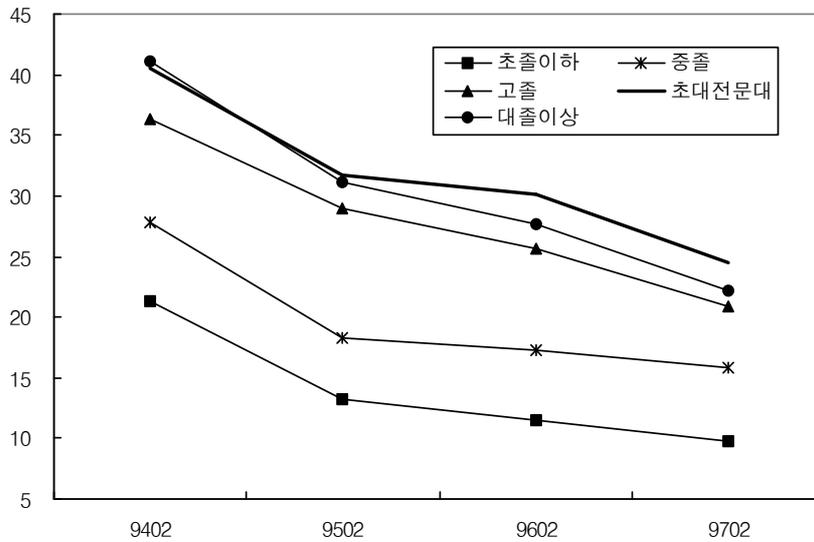
모든 교육수준별 집단에서 조사가 거듭될수록 이탈률이 감소하고 있어서 성과 연령 집단에서 나타난 특징이 교육정도별 집단에서도 그대로 나타나고 있다. 이러한 특징은 학력이 높을수록 보다 강하게 나타난다. 대졸이상의 교육수준을 가지고 있는 사람들을 보면 2차 조사에서 이탈률이 41.1%로 매우 높으나 3차 조사에서는 31.1%로, 4차 조사에서는 27.7%로, 그리고 5차 조사에서는 다시 22.2%로 크게 하락하고 있다. 즉 대졸이상 학력자의 5차 조사 이탈률은 2차 조사에 비해 18.9% 포인트로 크게 하락하고 있다. 이러한 이탈률 하락폭은 초졸이하의 11.7% 포인트, 중졸의 11.9% 포인트, 고졸의 15.6% 포인트, 그리고 초대·전문대의 16.1% 포인트 보다 더 크다.

이처럼 교육수준에 따라 이탈률에 큰 차이가 있고 조사가 거듭됨에 따른 이탈률 변화의 정도도 교육수준별로 다르다는 사실은 패널조사를 통해 얻는 교육수준별 통계치들의 신뢰성을 떨어뜨리게 된다.

<표3> 이탈자의 교육수준별·조사회수별 특성

		(단위 : 명, %)				
구	분	9302	9402	9502	9602	9702
표본 및 이탈자수	초졸이하	19,746	4,221	2,053	1,553	1,162
	중졸	13,256	3,679	1,758	1,347	1,026
	고졸	31,720	11,541	5,839	3,687	2,219
	초대·전문대	2,903	1,177	548	355	201
	대졸이상	10,840	4,453	1,986	1,218	707
이탈률	초졸이하		21.4	13.2	11.5	9.7
	중졸		27.8	18.4	17.2	15.9
	고졸		36.4	28.9	25.7	20.8
	초대·전문대		40.5	31.7	30.1	24.4
	대졸이상		41.1	31.1	27.7	22.2

자료 : 경제활동인구조사 원자료



<그림2> 교육정도별 · 조사회수별 특징

【혼인상태별 특성】

표본 이탈자들을 혼인상태별로 보면 배우자가 있는 유배우와 배우자와 사별한 경우는 이탈률이 낮고 이혼하였거나 미혼인 경우는 이탈률이 높다. 그리고 이혼한 경우를 제외하고는 조사가 계속됨에 따라 이탈률이 모두 하락하지만, 이혼한 사람들은 3차 조사 이후 조사가 거듭될수록 오히려 이탈률이 높아지는 특이한 모습을 보인다.

먼저 혼인상태별 이탈률 차이에 대해 살펴보자. <표4>와 <그림3>에서 보는 바와 같이 2차 조사에서 유배우와 사별의 경우 이탈률이 28.7%와 26.2%로 낮다. 그러나 미혼과 이혼의 경우는 각각 40.6%와 43.2%로 매우 높다. 이러한 현상은 이후의 계속된 조사에서도 마찬가지로 나타난다. 미혼인 경우는 대개 연령이 낮으므로 이들의 이탈률이 높다는 사실은 앞에서 연령이 낮을수록 이탈률이 높다는 사실과 밀접하게 관련되어 있는 것으로 판단된다. 이혼한 경우는 이혼에 대한 좋지 않은 사회적 인식과 이혼으로 인한 가정해체에 따른 표본 이탈 등이 이들의 이탈률을 높이는 주요 원인인 것으로 생각된다.

거듭된 조사에 따른 이탈률의 변화는 <그림3>에서 분명하게 볼 수 있다. 미혼과 유배우 및 사별의 경우에는 조사가 계속될수록 이탈률이 모두 하락하고 있으며 이탈률 하락의 정도는 유배우가 가장 크고 미혼이 그 다음이며 사별이 가장 작다. 반면 이혼한 사람들의 경우 2차 조사에서 43.2%였던 이탈률이 3차 조사에서는 31.2%로 하락하나 4차 조사에서는 35.5%로 다시 높아지며 5차 조사에서는 42.6%로 크게 상승하고 있다. 이는 성이나 연령, 그리고 교육정도에 따라 분류한 모든 집단에서 나타났던 현상과 상반되는 특징이다¹⁴⁾.

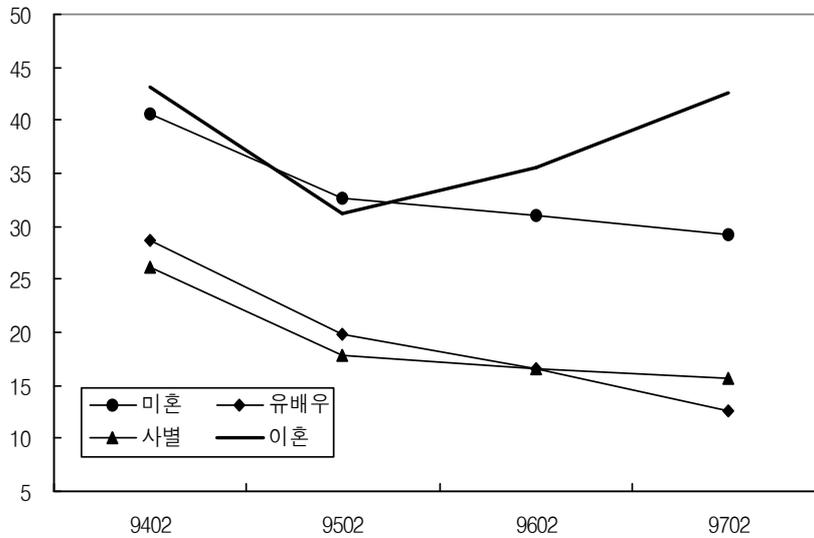
14) 이혼자들이 왜 이러한 구분되는 특징을 보이는가에 대해 언급하기 위해서는 별도의 분석이 필요하다.

<표4> 이탈자의 혼인상태 · 조사회수별 특성

(단위 : 명, %)

구 분		9302	9402	9502	9602	9702
표본 및 이탈자수	미 혼	22,008	8,928	4,265	2,729	1,775
	유 배 우	49,152	14,093	6,925	4,665	2,941
	사 별	6,504	1,704	852	655	513
	이 혼	801	346	142	111	86
이탈률	미 혼		40.6	32.6	31.0	29.2
	유 배 우		28.7	19.8	16.6	12.5
	사 별		26.2	17.8	16.6	15.6
	이 혼		43.2	31.2	35.5	42.6

자료 : 경제활동인구조사 원자료



<그림3> 혼인상태별 · 조사회수별 특성

이처럼 혼인상태에 따라 표본의 이탈률이 서로 다르고 패널조사가 몇 번 실시되었느냐에 따라 이탈률이 크게 차이가 나므로, 패널조사 결과를 사용한 혼인상태별 자료들의 신뢰성 역시 매우 낮아지게 된다. 특히 이혼한 경우 이탈률의 변화는 매우 급격하므로 이들에 대한 별도의 고려 없이 통계자료를 작성할 경우 그러한 통계치의 편의는 매우 클 것이다.

【가구주와의 관계별 특성】

표본 이탈자들을 가구주와의 관계로 구분하여 이들이 어떤 특징을 가지는지 살펴보자. <표5>와 <그림4>¹⁵⁾에서 보듯이 가구주를 비롯하여 배우자, 미혼자녀, 부모, 조부모, 그리고 기혼자녀 등 주로 가구주와 직계존속 및 미혼 직계비속의 관계를 가지는 사람들의 이탈률이 상대적으로 낮다. 반면 기혼자녀, 미혼 형제 자매, 기혼 형제 자매, 기타 친인척, 그리고 동거인의 경우 이탈률이 높다. 이들 집단은 일반적으로 단기간 동안만 가구주와 동일한 가구를 형성하며 가구가 분리될 가능성이 높은 집단이다. 따라서 이들의 이탈률은 주로 직계로 구성되는 가구원들에 비해 높을 것이다. 예를 들어 동거인의 경우 2차 조사에서 이탈률이 80%에 육박할 정도로 높은 것은 이러한 상황을 반영하고 있는 것으로 판단된다.

조사가 진행될수록 이탈률이 하락하는 특징은 대부분의 경우에서 나타난다. 이러한 특징은 표본의 대부분을 차지하는 가구주와 배우자에게서 두드러지며 부모와 미혼형제자매에서도 잘 나타난다. 동거인에서는 이러한 특징이 극단적으로 나타난다¹⁶⁾. 한편 미혼자녀의 경우 이탈률이 조사회수와 특별한 관계를 가지고 있다고 보기 어렵다. 2차 조사에서 33.0%였던 이들의 이탈률이 3차 조사에서 30% 수준으로 하락하고 있으나 4차와 5차 조사에서 이탈률은 30.9%에서 일정하게 유지되고 있다. 이는 결혼에 따른 분가 등으로 인한 표본 이탈이 일정한 정도로 존재함을 의미하는 것으로 해석할 수 있겠다. 미혼 형제 자매의 경우에도 조사가 계속됨에 따른 이탈률의 하락이 높지 않아 미혼자녀의 경우와 닮은 모습을 보인다.

가구주와 어떤 관계에 있느냐에 따라 표본 이탈률이 크게 다르고 조사가 몇 차례 실시되었느냐에 따라 이들 사이의 이탈률의 변화가 급격하다는 점은 패널조사를 통해 가구주와의 관계별로 자료를 작성할 경우 그러한 자료의 신뢰성을 크게 저해하게 될 것임을 의미한다. 가구주와의 관계에 따른 이탈률 격차 및 조사회수에 따른 이탈률 변화의 차이는 성이나 연령 등 이제까지 보았던 경우들 보다 더 크다. 따라서 신뢰성 저하의 정도 역시 더욱 클 것이므로 패널조사에서 이에 대한 특별한 주의가 필요하다.

15) 표와 그림 및 본문에서 ‘부모’에는 장인·장모가 포함되며, ‘기혼자녀’와 ‘기혼형제자매’에는 그 가족도 포함된다.

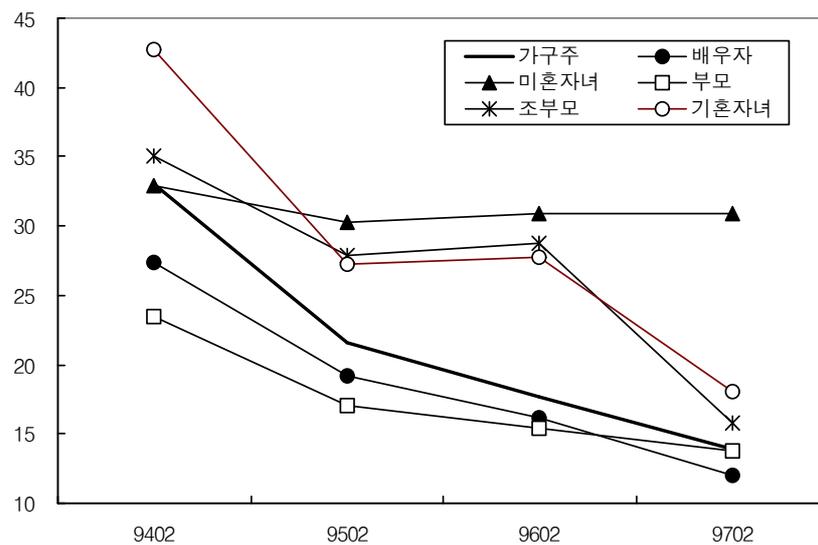
16) 동거인은 대개 소속된 가구에 대한 정착도가 낮아 이미 2차 조사에서 80%에 가까운 표본이 이탈하여 ‘빠져나갈 사람들은 이미 대부분 빠져나간’ 것이 이탈률 하락의 주된 원인으로 생각된다.

<표5> 이탈자의 가구주와의 관계별 · 조사회수별 특징

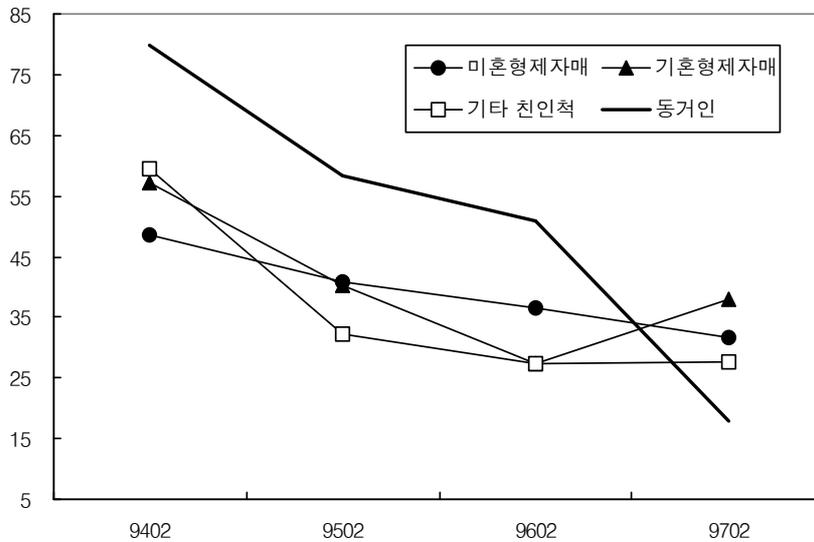
(단위 : 명, %)

구분	9302	9402	9502	9602	9702	
표본 및 이탈자수	가구주	30,591	10,121	4,429	2,840	1,839
	배우자	22,899	6,279	3,197	2,177	1,355
	미혼자녀	14,773	4,871	2,994	2,137	1,472
	기혼자녀	1,157	495	180	134	63
	부모	5,273	1,237	686	515	390
	조부모	171	60	31	23	9
	미혼형제자매	2,420	1,175	507	269	149
	기혼형제자매	356	204	61	25	25
	기타 친인척	146	87	19	11	8
	동거인	679	542	80	29	5
이탈률	가구주		33.1	21.6	17.7	13.9
	배우자		27.4	19.2	16.2	12.0
	미혼자녀		33.0	30.2	30.9	30.9
	기혼자녀		42.8	27.2	27.8	18.1
	부모		23.5	17.0	15.4	13.8
	조부모		35.1	27.9	28.8	15.8
	미혼형제자매		48.6	40.7	36.4	31.8
	기혼형제자매		57.3	40.1	27.5	37.9
	기타 친인척		59.6	32.2	27.5	27.6
	동거인		79.8	58.4	50.9	17.9

자료 : 경제활동인구조사 원자료



<그림4> 가구주와의 관계별 · 조사회수별 이탈률1



<그림5> 가구주와의 관계별 · 조사회수별 이탈률2

【경제활동상태별 특성】

이제 표본에서 이탈한 사람들을 전년도 경제활동상태에 따라 분류해 보자. <표6>과 <그림6>은 이를 나타낸 것이다. 표와 그림에서 보는 바와 같이 전년도 조사에서 실업상태에 있었던 사람들의 이탈률이 가장 높다. 그리고 전년도에 어떤 경제활동상태에 있었느냐와 무관하게 조사가 계속될수록 이탈률이 감소하고 있다.

실업자의 이탈률은 2차 조사에서 41.8%로 높지만 취업자는 33.9%, 그리고 비경제활동인구는 29.0%로 상대적으로 낮은 이탈률을 보인다. 실업에 따른 신분의 불안정과 실업에 대한 좋지 않은 사회의 인식 등은 이들의 이탈률을 높게 하는 주요 원인으로 작용할 것이다. 반면 취업자와 비경제활동인구의 이탈률은 2차 조사에서 4.9% 포인트의 차이를 보이지만 4차 및 5차 조사에서는 양자의 이탈률에 큰 차이가 없다.

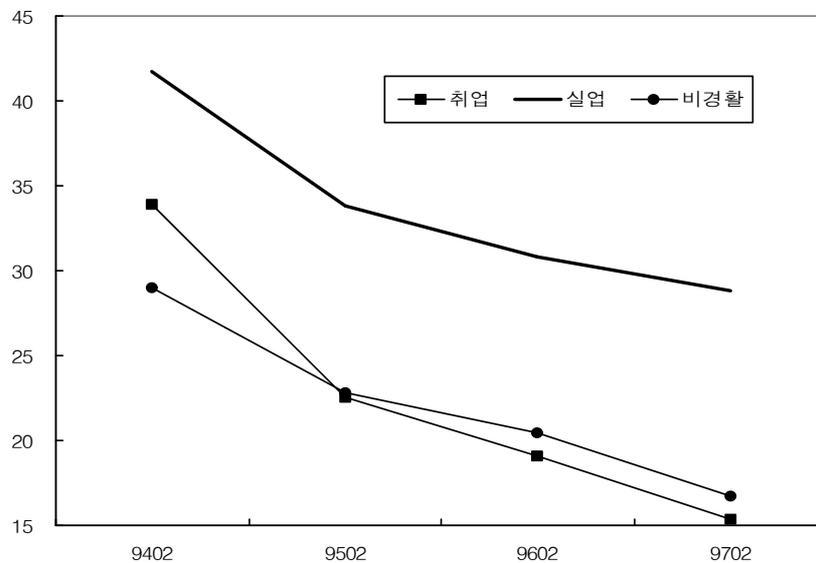
한편 조사가 진행됨에 따른 이탈률의 하락은 모든 경제활동상태별 집단에서 공통적으로 나타나지만 취업자에게서 두드러진다. 2차 조사에서 33.9%였던 취업자의 이탈률이 3차 조사에서 22.5%, 4차 조사에서 19.1%로 하락하였으며 5차 조사에서는 15.4%로 급속히 감소하고 있다. 즉 2차 조사에 비해 5차 조사에서 이탈률이 18.5% 포인트 하락하였으며 이는 실업자의 13.0% 포인트나 비경제활동인구의 12.2% 포인트 보다 크다.

<표6> 이탈자의 경제활동상태별 · 조사회수별 특성

(단위 : 명, %)

구분		9302	9402	9502	9602	9702
표본수	취업	42,881	30,265	24,008	19,556	16,681
	실업	1,530	829	506	330	339
	비경활	34,054	22,300	16,696	13,164	10,715
이탈자수	취업		14,549	6,810	4,582	3,013
	실업		639	280	156	95
	비경활		9,883	5,094	3,422	2,207
이탈률	취업		33.9	22.5	19.1	15.4
	실업		41.8	33.8	30.8	28.8
	비경활		29.0	22.8	20.5	16.8

자료 : 경제활동인구조사 원자료



<그림6> 경제활동상태별 · 조사회수별 이탈률

실업상태에 있었던 사람들의 이탈률이 취업이나 비경제활동상태에 있었던 사람들 보다 더 크고 이들 간의 이탈률 차이도 2차, 3차 조사에서 크게 달라짐에 따라 패널조사를 통해 작성한 실업관련 각종 통계의 신뢰성에 대해 의문이 제기될 수 있다. 예를 들어 가구주와의 관계에 따라 그리고 이들에 대한 조사가 몇 번 실시되었느냐에 따라 이탈률에 큰 차이가 있으므로 가구주와의 관계별로 실업률을 작성할 경우 패널조사 결과는 매우 큰 편의를 가질 것이다.

【종사상의 지위별 특성】

다음으로 전년도에 취업상태에 있었던 사람들 중 표본에서 이탈한 사람들이 전년도에 종사하던 지위별로 어떤 구분되는 특징을 보이는지 살펴보자. <표7>과 <그림7>, 그리고 <그림8>은 이를 정리한 것이다. 표와 그림에서 보는 바와 같이 임금근로자의 이탈률이 비임금근로자의 경우 보다 매우 높다. 임금근로자 가운데에는 임시근로자가, 그리고 비임금근로자 가운데에는 고용주의 이탈률이 가장 높다. 보다 구체적으로 살펴보자.

먼저 임금근로자와 비임금근로자를 비교하여 보면, <표7>과 <그림7>에서 보는 바와 같이 2차 조사에서 비임금근로자의 이탈률은 26.5%이나 임금근로자의 경우는 이 보다 11.8% 포인트나 높은 38.4%이다. 임금근로자 가운데 상용근로자만 보더라도 비임금근로자인 고용주, 자영주, 그리고 무급가족종사자의 경우 보다 이탈률이 더 높다. 이러한 이탈률 격차는 거듭된 조사에서도 마찬가지로 나타나며 그 격차도 12% 포인트 정도로 안정되어 있다. 왜 임금근로자의 이탈률이 비임금근로자들 보다 높은가에 대해 분명한 언급을 하기 위해서는 이에 대한 별도의 엄밀한 연구가 뒤따라야 할 것이다.

<표7>과 <그림8>에서 보듯이 임금근로자 가운데에도 임시근로자의 이탈률이 가장 높으며 일용근로자의 이탈률이 가장 낮다. 2차 조사에서 일용근로자의 이탈률은 상용근로자의 경우와 동일한 37.2%이며 임시근로자의 이탈률은 41.9% 이다. 3차 조사 및 그 이후부터는 상용근로자의 이탈률이 일용근로자보다 높게 나타나 일용, 상용, 임시 근로자의 순으로 이탈률이 높다. 이탈률이 가장 높은 임시근로자의 경우, 조사가 계속됨에 따른 이탈률 하락 정도도 타 임금근로자나 비임금근로자들에 비해 미미하다. 2차 조사에 비해 5차 조사에서 이탈률 하락폭은 일용근로자가 18.8% 포인트로 가장 크고 그 다음이 상용근로자로서 14.9% 포인트인 반면, 임시근로자는 9.3% 포인트에 불과하다.

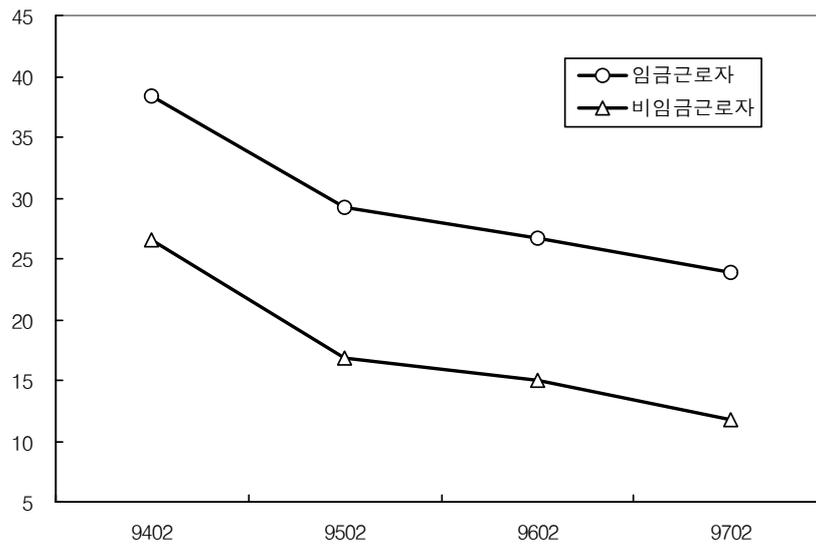
비임금근로자의 경우를 보면 이들 중 가장 큰 비중을 차지하는 자영자의 이탈률이 24.8%로서 무급가족종사자의 25.8%와 함께 낮고 고용주의 이탈률이 32.7%로 높다. 3차 및 그 이후의 조사에서 자영자와 무급가족종사자의 이탈률은 거의 차이를 보이지 않고 있으며 고용주 이탈률과의 격차는 더욱 커지고 있다. 즉 고용주는 이탈률이 높을 뿐만 아니라 거듭된 조사에서 이탈률이 하락하는 정도도 크지 않다. 고용주의 이탈률은 2차 조사에 비해 5차 조사에서 13.3% 포인트 하락하였는데, 이는 자영자의 14.4% 포인트 및 무급가족종사자의 15.7% 포인트 보다 작다.

<표7> 이탈자의 종사상의 지위별 · 조사회수별 특성

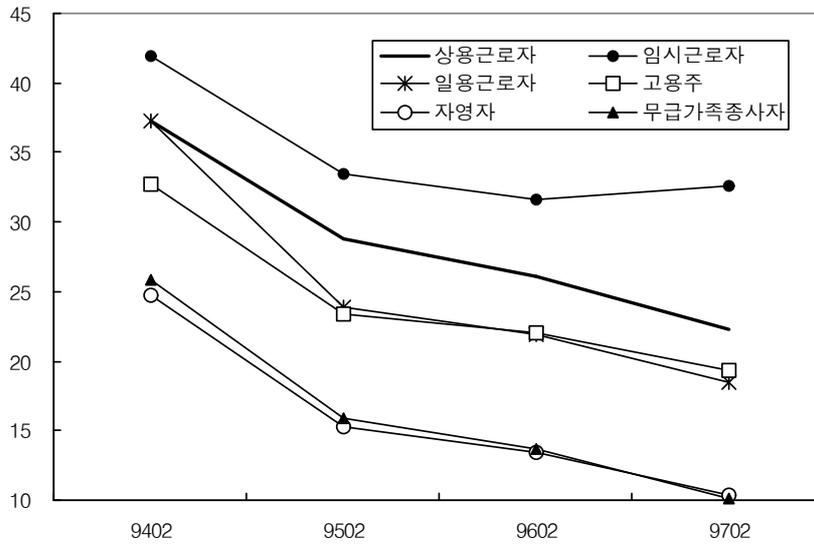
(단위 : 명, %)

구 분		9302	9402	9502	9602	9702
표본 및 이탈자수	임금근로자	26,593	10,216	4,790	3,097	2,024
	상용근로자	15,989	5,950	2,896	1,867	1,175
	임시근로자	6,850	2,868	1,331	838	591
	일용근로자	3,754	1,398	563	392	258
	비임금근로자	16,288	4,333	2,020	1,485	989
	고 용 주	3,198	1,047	503	364	249
	자 영 자	9,023	2,235	1,037	775	518
	무급가족 종 사 자	4,067	1,051	480	346	222
이탈률	임금근로자		38.4	29.2	26.7	23.8
	상용근로자		37.2	28.8	26.1	22.3
	임시근로자		41.9	33.4	31.6	32.6
	일용근로자		37.2	23.9	21.9	18.4
	비임금근로자		26.6	16.9	14.9	11.7
	고 용 주		32.7	23.4	22.1	19.4
	자 영 자		24.8	15.3	13.5	10.4
	무급가족 종 사 자		25.8	15.9	13.6	10.1

자료 : 경제활동인구조사 원자료



<그림7> 종사상지위별 · 조사회수별 이탈률1



<그림8> 종사상지위별 · 조사회수별 이탈률2

3. 대우패널조사에서 표본 이탈자의 특성

1) 표본 이탈과 패널자료의 대표성

【표본의 수와 표본 이탈률】

【성별 특성】

2) 표본 가구의 이탈과 그 특성

IV. 결론

< 참고문헌 >

- 김대일, “실업장기화의 효과분석,” 한국의 노동경제 — 쟁점과 전망 —, 경문사, 1999, 4.
- 남재량, 우리나라의 실업률 추세변화에 관한 연구, 서울대학교 박사학위 논문, 1997. 8.
- 남재량·류근관, “우리나라 여성 노동력 상태의 동태적 특성에 관한 연구,” 한국사회과학, 제21권 3호, 2000.
- 남재량·이창용, “한국의 실업률 추세변화에 관한 연구,” 경제학연구 제46집 제2호, 1998.6
- 대우경제연구소, 『한국 가구 경제활동 연구』, 1994. 7.
- _____, 『한국인의 경제활동』, 1995, 1996, 1997.
- 류근관, 지난 30여년간의 한국노동시장의 수급실태조사 : 노동공급과 임금의 추계를 중심으로, *****
- 신동균, 미국 패널데이터의 현황과 시사점 — PSID, NLSY, KHPS, KLIPS를 중심으로 —, 한국노동패널연구 97-02, 1998. 4.
- Beckett, S., W. Gould, L. Lillard, and F. Welch, “The Panel Study of Income Dynamics after Fourteen Years: An Evaluation,” *Journal of Labor Economics*, Oct. 1988.
- Brown, C., D. Greg, and F. Stafford, “Data Watch: The Panel Study of Income Dynamics,” *Journal of Economic Perspectives*, Spring 1996.
- Center for Human Resource Research, *The National Longitudinal Surveys: User’s Guide*, The Ohio State University, Columbus, Ohio, 1993.
- _____, *The National Longitudinal Surveys: Handbook*, The Ohio State University, Columbus, Ohio, 1995.
- Duncan, G., and D. Hill, “Assessing the Quality of Household Panel Survey Data : The Case of the PSID,” *Journal of Business and Economic Statistics*, Oct. 1989.
- Fahy, T.W., *The National Longitudinal Surveys of Labor Market Experience An Annotated Bibliography of Research*, The Ohio State University, 1995.
- Fitzgerald, J., P. Gottschalk, and R. Motiff, “A Study of Sample Attrition in the Michigan Panel Study of Income Dynamics,” Prepared for the Conference on Nonresponse in Panels, Washington D.C., Feb. 1994.
- Institute for Social Research, *Proposal to the National Science Foundation*, Ann Arbor, Michigan, 1996.
- Martha S. Hill, *Panel Study of Income Dynamics: A User’s Guide*, Newbury Park, Ca. : Sage Publications, 1992
- Ryu, Keunkwan, “Analysis of Continuous Time Proportional Hazard Model Using Discrete Duration Data,” *Econometric Reviews*, 1995.
- _____, “Group Duration Analysis of Proportional Hazard Model: Minimum Chi-square Estimators and Specification Tests,” *JASA*, 1994.
- _____, “Structural Duration Analysis of Management Data,” *Journal of Econometrics*, 1993.
- Statistics Canada, *Dealing with the Seam Problem for the Survey of Labour and Income Dynamics*, 1992.
- _____, *SLID: Microdata User’s Guide*, 1994.

<부록> 한국노동연구원 패널 사용시 각종 변수의 정의

1. 가구 자료

가구근로소득 유무 : 근로소득 있음, V01440=1
가구근로소득 : V01441 (V01440=1 인 경우)

2. 개인 자료

(1) 경제활동 상태

취업자 : 설문지 종류, P01573=1 또는 2
실업자 : 설문지 종류, P01573=3 &
지난 1주(또는 1개월)간 구직여부, P01203=1 (P01204=1) &
취업 가능성 여부, P01205=1
비경제활동 : 설문지 종류, P01573=3 &
구직여부, P01203≠1 & P01204≠1 또는
취업 가능성 여부, P01203=1 또는 P01204=1, 그러나 P01205≠1

(2) 취업자 변수(설문지 종류, P01573=1 또는 2, 군인제외, P01502≠2)

(종사상 지위)

상용근로자 : 설문지 종류, P01573=1 & 종사상 지위, P01026=1
임시직 : 설문지 종류, P01573=1 & 종사상 지위, P01026=2
일용직 : 설문지 종류, P01573=1 & 종사상 지위, P01026=3 또는 9(모름)
고용주 : 설문지 종류, P01573=2 & 피고용자 유무, P01136=1 & P01572≠3
자영업주 : 설문지 종류, P01573=2 & 피고용자 유무, P01136=2 & P01572≠3
무급가족 : 설문지 종류, P01573=2 & 1주간 취업상태, P01572≠3

(기업규모) 자영업주/무급가족 종사자 제외

1~4인 : 설문지 종류, P01573=1 & 기업규모, P01016=2 & P01018=1, 또는
설문지 종류, P01573=1 & 기업규모, P01016=1 & $1 \leq P01017 \leq 4$
설문지 종류, P01573=2 & 기업규모, P01137=2 & P01139=1, 또는
설문지 종류, P01573=2 & 기업규모, P01137=1 & $1 \leq P01138 \leq 4$
5~9인 : 설문지 종류, P01573=1 & 기업규모, P01016=2 & P01018=2, 또는
설문지 종류, P01573=1 & 기업규모, P01016=1 & $5 \leq P01017 \leq 9$
설문지 종류, P01573=2 & 기업규모, P01137=2 & P01139=2, 또는
설문지 종류, P01573=2 & 기업규모, P01137=1 & $5 \leq P01138 \leq 9$
10+인 : 설문지 종류, P01573=1 & 기업규모, P01016=2 & P01018≥3, 또는
설문지 종류, P01573=1 & 기업규모, P01016=1 & $10 \leq P01017$
설문지 종류, P01573=2 & 기업규모, P01137=2 & P01139≥3, 또는
설문지 종류, P01573=2 & 기업규모, P01137=1 & $10 \leq P01138$

(근로 특성) 임금근로자(설문지 종류, P01573=1 & (P01015<=4 또는 P01015>=10))

월 근로일수 : P01049(≠99) if P01036=1(근로시간 규칙성) 또는
(P01048(≠99)+P01042(≠99))/2 if P01036=2 또는 3(불규칙적)

주당근로시간 : P01054(≠99) if P01053==1(정규시간 유무) 또는
P01060(≠99) if P01053==2(비정규적인 경우)

직장시작일 : (년도) P01030(≠9999)-1900 if P01029==1(정규직?) 또는
P01034(≠9999)-1900 if P01029==2(비정규직)
(월) P01031(≠99) if P01029==1(정규직) 또는
P01035(≠99) if P01029==2(비정규직)

근속개월 : (1998-직장시작년도)*12 + (P01541(=면접월)-직장시작월)

(임금소득) 임금근로자(설문지 종류, P01573=1 & (P01015<=4 또는 P01015>=10))

세후 임금 : P01091(≠9999)

(구직기간) 실업자(설문지 종류, P01573=3 & 구직여부, P01203=1 (또는 P01204=1)
& 취업 가능성 여부, P01205=1)

구직기간 : (년수) P01230(≠99), 99인 경우에는 0으로 처리
(월수) P01231(≠99), 99인 경우에는 0으로 처리
(일수) P01232(≠9), 9인 경우에는 0으로 처리

구직개월 : = 구직년수*12 + 구직월수 + 구직일수*7/31