

# [패널자료 품질개선 연구(VIII)]

[www.kli.re.kr](http://www.kli.re.kr)

김유빈 · 박인호 · 이지은 · 신선옥  
이혜정 · 정현상 · 최세림 · 박경원



## 목 차

요 약 .....	i
제1장 서 론 .....	(김유빈) ..... 1
제2장 제6차 사업체패널조사 가중치 산출 연구 ....	(박인호) ..... 6
제1절 서 론 .....	6
제2절 기존 표본설계 및 응답현황 .....	8
1. 기존 표본설계 .....	8
2. 기존 조사 응답현황 .....	11
제3절 6차년도 조사의 표본추가를 위한 표본설계 및 응답 현황 .....	15
1. 표본추가를 위한 이중추출틀 방식의 표본설계 .....	15
2. 응답현황 .....	21
제4절 사업체패널조사 가중치 산정 .....	26
1. 제6차 조사 종단면가중치 .....	26
2. 제6차 조사 횡단면가중치 .....	29
제3장 웹을 활용한 조사 효과에 대한 실증연구 II: 노동패널 CAWI 1차, 2차 실험 연구 결과 분석 ..... (이지은 · 이해정 · 신선옥 · 정현상) .....	40
제1절 서 론 .....	40
제2절 노동패널 CAWI 2차 실험 설계 개요 .....	42
제3절 CAWI 1차 2차 조사에 관한 실증분석 .....	44

1. 표본유지율 .....	44
2. 가구 및 개인의 특성 차이 .....	50
3. 분포의 차이 .....	59
4. 항목무응답률 .....	74
5. 설문유형 변경 실험 분석 .....	78
6. 직산업 코딩 검토 .....	81
제4절 소 결 .....	88

## 제4장 출산력변수 형성방법 및 문제점 진단 ..... (최세림) ..... 98

제1절 서 론 .....	98
제2절 출산력 관련 설문문항 및 변수 검토 .....	99
제3절 출산력 패널자료 구축방법 검토 .....	105
1. 총 자녀수 패널 변수 생성방법의 예 .....	107
2. 남성의 총 자녀수 변수 생성방법의 예 .....	112
3. 기타 총 자녀수 패널 변수 생성방법 검토 .....	114
제4절 노동패널로 구축할 수 있는 출산력 자료의 기초통계 .....	117
제5절 소 결 .....	120

## 제5장 사업체패널조사의 경영의사결정 참여 설문문항들의

### 효과적 활용 탐색 ..... (박경원) ..... 123

제1절 서 론 .....	123
제2절 경영의사결정 참여 설문문항의 분석 개요 .....	125
1. 신뢰도와 타당도 .....	125
2. 경영의사결정 참여 설문문항을 분석한 국내 연구 .....	128
제3절 데이터 및 변수의 측정 .....	129
1. 데이터 .....	129
2. 변수의 측정 .....	130

제4절 경영의사결정 참여수준의 계량적 분석 .....	136
1. 경영의사결정 참여 설문문항의 신뢰도 분석 .....	136
2. 경영의사결정 참여수준의 상관관계 분석 .....	138
3. 경영의사결정 참여수준의 결정요인 분석 .....	143
4. 경영의사결정 참여수준과 사업체 성과 분석 .....	145
제5절 소 결 .....	152
참고문헌 .....	157

## 표 목 차

<표 1- 1> 패널자료 품질개선 연구(I~VII) 세부 연구주제 .....	2
<표 2- 1> 공공부문을 제외한 사업체패널 1차 조사의 표본수 및 추출률 .....	9
<표 2- 2> 2006년 사업체패널조사 조사 현황 .....	10
<표 2- 3> 2006년 사업체패널조사(1차 조사)에 대한 표본 사업체의 응답률 및 표본 부적격률 .....	11
<표 2- 4> 사업체패널조사 1~5차 패널사업체 추적조사 응답현황 및 패널유지율 .....	13
<표 2- 5> 사업체패널조사 1~5차 전 차수 대비 표본유지율 .....	13
<표 2- 6> 사업체패널조사의 1~5차의 전체 사업체에 대한 응답유형 현황 .....	14
<표 2- 7> 업종분류 .....	17
<표 2- 8> 조사모집단의 산업×규모별 사업체 수 .....	18
<표 2- 9> 1차 신규표본 추가분 .....	19
<표 2-10> 2차 신규표본 추가분 .....	20
<표 2-11> 전체 신규표본 추가 .....	20
<표 2-12> 제6차년도 표본종류별 응답현황 .....	21
<표 2-13> 사업체패널조사 1~6차 전체 사업체에 대한 응답유형 현황 .....	22
<표 2-14> 제6차년도 표본종류별 횡단면조사 기준 응답현황 .....	23
<표 2-15> 제6차년도 기존표본 횡단면조사 산출 대상 표본사업체 현황 .....	23
<표 2-16> 제6차년도 기존표본 횡단면조사 기준 최종 응답현황 ....	24
<표 2-17> 제6차년도 신규표본 횡단면조사 산출 대상 표본사업체	

현황 .....	25
<표 2-18> 제6차년도 신규표본 횡단면조사 산출 대상 응답사업체 현황 .....	25
<표 2-19> 1~5차 사업체패널조사의 전체 사업체 응답유형 현황 ...	28
<표 2-20> 사업체패널조사의 5~6차 패널가중치 분포 .....	28
<표 2-21> 사업체패널조사의 사후층 크기 .....	33
<표 2-22> 이중틀 표본종류와 산업분류별 응답사업체 현황 .....	35
<표 2-23> PML 횡단면가중치 산출 구성요소 .....	36
<표 2-24> 산업분류 및 노동조합 여부에 따른 사업체 분포 .....	38
<표 2-25> 사업체규모 및 노동조합 여부에 따른 분포 .....	38
<표 2-26> 사업체패널조사의 제6차 횡단면 가중치 분포 .....	39
<표 3- 1> 조사도구별 패널유지율 차이 .....	45
<표 3- 2> 2차 조사의 개인 식별 변수(pid) 매칭 오류 .....	46
<표 3- 3> 2차 조사의 개인 식별 변수(pid) 매칭 오류 유형 .....	46
<표 3- 4> 조사도구별 패널 유지 탈락별 가구원수 차이 검증 .....	48
<표 3- 5> 조사도구별 패널 유지 탈락별 지역 차이1 .....	48
<표 3- 6> 조사도구별 패널 유지 탈락별 지역 차이2 .....	48
<표 3- 7> 조사도구별 패널 유지 탈락별 성별 차이 .....	49
<표 3- 8> 조사도구별 패널 유지 탈락별 연령 차이 .....	50
<표 3- 9> 조사도구별 패널 유지 탈락별 학력 차이 .....	50
<표 3-10> 조사도구별 패널 유지 탈락별 동거 여부 차이 .....	50
<표 3-11> 가구 근로소득액 응답률(CAPI vs CAWI) .....	51
<표 3-12> 가구소득 응답률의 집단(CAPI vs CAWI) 간 차이 .....	52
<표 3-13> 금융자산 응답률의 집단(CAPI vs CAWI) 간 차이 .....	53
<표 3-14> 지난 한 달 소득 응답률의 집단(CAPI vs CAWI) 간 차이 .....	53
<표 3-15> 월평균 저축액 응답률의 집단(CAPI vs CAWI) 간 차이 .....	54
<표 3-16> 가구 생활비 응답률의 집단(CAPI vs CAWI) 간 차이 ...	55

<표 3-17> 주거지 정보 응답률의 집단(CAPI vs CAWI) 간 차이 ..	55
<표 3-18> 임금근로자 비중 조사방법에 따른 차이 비교 .....	56
<표 3-19> 종업원수 응답률의 조사방법에 따른 차이 비교 .....	56
<표 3-20> 흡연 및 음주 응답비중의 집단 간 차이 .....	57
<표 3-21> 민감 문항 응답률의 집단 간 차이 .....	58
<표 3-22> 입주·주거형태 및 가구원수 차이 검증 .....	59
<표 3-23> 소득변수 차이 검증 .....	60
<표 3-24> 소득변수에 대한 이상치 개수 .....	61
<표 3-25> 소득변수 차이 검증:이상치 제거 후 분석 .....	62
<표 3-26> 이전소득액 및 이전소득액의 세부 항목별 응답가구수 ...	63
<표 3-27> 금융자산 및 기타변수 차이 검증 .....	63
<표 3-28> 금융자산 및 기타변수에 대한 이상치 개수 .....	64
<표 3-29> 금융자산 및 기타변수 차이 검증:이상치 제거 후 분석 ...	64
<표 3-30> 직장만족도 응답 차이 검증 .....	66
<표 3-31> 생활만족도 차이 검증 .....	67
<표 3-32> 건강상태 및 장애 응답 차이 검증 .....	68
<표 3-33> 흡연 및 음주 변수의 차이 검증 .....	69
<표 3-34> 종사상 지위 분포의 조사방법에 따른 차이 비교 .....	70
<표 3-35> 임금근로자 중 비정규직 근로자 비중의 조사방법에 따른 차이 비교 .....	71
<표 3-36> 비정규직 근로형태별 조사방법에 따른 차이 비교 (통계기준) .....	72
<표 3-37> 평균근로시간의 조사방법에 따른 차이 비교 .....	73
<표 3-38> 평균소득 유무의 조사방법에 따른 차이 비교 .....	73
<표 3-39> 평균소득의 조사방법에 따른 차이 비교 .....	74
<표 3-40> 재학 여부, 혼인상태 무응답률(CAPI vs CAWI) .....	75
<표 3-41> 작년 한 해 근로소득 무응답률의 집단 간 차이 .....	76
<표 3-42> 일자리 시작 시기(연/월/일) 무응답률의 집단 간 비교 ...	77
<표 3-43> 임금근로자 중 비정규직 여부와 종업원수 무응답률의 조사방법에 따른 차이 비교 .....	77



<표 3-44> 평균소득 무응답률의 조사방법에 따른 차이 비교 .....	78
<표 3-45> 사회보험 수급종류별 응답률의 집단 간 차이 분석 .....	80
<표 3-46> 사회보험 항목배열 변경에 따른 차이 .....	81
<표 3-47> 산업코드의 코더 간 불일치율 .....	82
<표 3-48> 직업코드의 코더 간 불일치율 .....	82
<표 3-49> 코더별 임금근로자의 산업코드 분포 .....	84
<표 3-50> 코더별 임금근로자의 직업코드 분포 .....	85
<표 3-51> 코더별 비임금근로자의 산업코드 분포 .....	86
<표 3-52> 코더별 비임금근로자의 직업코드 분포 .....	87
<표 3-53> 조사 성공 관련 1차 조사결과 요약 .....	89
<표 3-54> 2차 패널 유지 결과 요약 .....	90
<표 3-55> 1차 2차 균형패널 조사결과 요약(액수 관련 변수 응답률) .....	91
<표 3-56> 1차 2차 균형패널 조사결과 요약(액수 관련 변수 분포- 평균 비교) .....	91
<표 3-57> 1차 2차 균형패널 조사결과 요약(일자리, 직업훈련) .....	92
<표 3-58> 1차 2차 균형패널 조사결과 요약(민감문항 응답률) .....	92
<표 3-59> 1차 2차 균형패널 조사결과 요약(직장만족도 응답 차이) .....	93
<표 3-60> 1차 2차 균형패널 조사결과 요약(생활만족도) .....	94
<표 3-61> 1차 2차 균형패널 조사결과 요약(건강상태 및 장애) .....	95
<표 3-62> 1차 2차 균형패널 조사결과 요약(일자리) .....	95
<표 3-63> 1차 2차 균형패널 조사결과 요약(직산업 코딩 불일치율) .....	96
<표 4- 1> 출산력 정보 구축을 위해 사용할 수 있는 변수 .....	101
<표 4- 2> 출산력 신규 응답자의 자녀수 문항 남녀 비교(p_9074+p_9075) .....	102
<표 4- 3> 2016년도 미취학아동 가구원수 비교 .....	104
<표 4- 4> 사교육 문항과 개인 출산경험 문항을 활용하여 출산력	

데이터 구축방법 예시 .....	110
<표 4- 5> 가구주와의 관계 코드 .....	113
<표 4- 6> 사교육 문항을 이용하여 생성한 신규출생자녀수 변수 ...	118
<표 4- 7> 신규가구원 편입이유변수를 사용하여 생성한 신규출생자녀수 변수 .....	119
<표 4- 8> 다양한 방법으로 생성한 자녀수 관련 변수들의 기초 통계 .....	120
<표 5- 1> 노동조합/노사협의회의 경영의사결정 참여 설문문항 .....	131
<표 5- 2> 경영의사결정 참여 설문문항에 대한 노사 응답의 신뢰도(2013년) .....	137
<표 5- 3> 경영의사결정 참여 설문문항의 내적 일관도 .....	137
<표 5- 4> 경영의사결정 참여수준의 평균과 표준편차 .....	139
<표 5- 5> 경영의사결정 참여수준과 Correlates .....	140
<표 5- 6> 패널고정효과 모형으로 본 경영의사결정 참여수준의 결정요인 .....	144
<표 5- 7> 패널고정효과 모형으로 본 참여수준의 노사관계 분위기에 대한 영향 .....	147
<표 5- 8> 패널고정효과 모형으로 본 참여수준의 고용안정성에 대한 영향 .....	148
<표 5- 9> 패널고정효과 모형으로 본 참여수준의 노동생산성에 대한 영향 .....	150
<표 5-10> 패널고정효과 모형으로 본 참여수준의 수익성에 대한 영향 .....	151

## 그림목차

[그림 2-1] 목표모집단 및 표본 분해도 .....	16
[그림 3-1] 개인 식별 변수(pid) 매칭 오류 유형 .....	46
[그림 3-2] 설문항목 방식 .....	79
[그림 4-1] mom_pid 생성방법 코드 일부 발췌 .....	108
[그림 4-2] 생성된 데이터 형태 예시 .....	109
[그림 5-1] 타당도 검증을 위한 실증모형 .....	126
[그림 5-2] 경영의사결정 참여의 단순합 분포(2005~2015년) .....	138



## 요 약

### 1. 서 론

본서는 「패널자료 품질개선 연구 시리즈」의 8차 결과물이다. 본 연구는 ‘제6차 사업체패널조사 가중치 산출 연구’, ‘웹을 활용한 조사 효과에 대한 실증연구 II’, ‘출산력변수 형성방법 및 문제점 진단’, ‘사업체패널조사의 경영의사결정 참여 설문문항들의 효과적 활용 탐색’에 관한 내용을 담고 있다.

제2장에서는 추가표본 추출로 변화된 표본구성을 반영하고자 적절한 표본가중치의 산출을 위한 연구를 수행하였다. 제6차 사업체패널조사를 위한 중단면가중치에 대해 논의하고, 개별적 횡단면가중치와 이중추출틀 접근에 기초한 통합표본의 의사최대가능도 횡단면가중치의 산출과정 및 결과를 논의한다.

제3장 ‘웹을 활용한 조사 효과에 대한 실증연구 II’는 패널자료 품질개선연구(VII)의 웹조사 연구에 이은 2년차 연구 프로젝트로, 표본 이탈에 CAPI- CAWI 조사그룹 간 이질성이 자료의 신뢰성 및 대표성에 미치는 영향을 분석하였다. 1차 실험조사와 2차 실험 조사의 결과를 비교 분석하고, CAWI가 대면조사인 CAPI의 한계를 극복할 수 있는 대안으로 가능한 것인가를 검토하는 데 목적을 둔다.

제4장은 주요 인구통계학적 변수로 꼽히는 출산 관련 변수에 주목하고, 출산력 구축 방안 및 이에 대응하여 발생하는 문제점을 파악하기 위한 연구를 수행하였다. 출산 관련 변수는 자녀의 출생이 여성 근로자들의 경력단절 및 임금에 미치는 영향, 일·가정 양립정책과 출산 영향 등 주요 연구에 있어 필수적 요소로 자리 잡고 있는 만큼, 본 연구에서는 출산력 구축방법을 정리하고, 출산력변수 생성과정에

서 발생하는 여러 문제점들을 논의하는 것을 목적으로 한다.

## 2. 제6차 사업체패널조사 가중치 산출 연구

제6차 사업체패널조사는 현재, 2006년 이후 지속된 기존패널과 2016년 표본 추가된 패널이 혼재한 분리패널(split panel)의 형태를 띠고 있다. 본 연구는 이러한 표본구성을 반영하기 위해 이중추출틀(dual frame) 접근하에 의사최대가능도(擬似最大可能度, pseudo-maximum likelihood, 이후 PML)이론을 이용한 표본가중치를 개발하였다. 이중추출틀 구조하에서 일반적으로 표본가중치가 반영해야 할 추출확률, 무응답조정, 이상치 가중값 조정, 모집단 보조정보의 반영 등이 복합적으로 반영할 수 있는 적절한 가중치 산출방법을 적용하는 것이 본 연구의 목적이다.

사업체패널조사는 신규표본 추가에 있어, 기존 패널모집단과 함께 표본추가 시점을 기준으로 횡단면모집단의 특성을 모두 반영할 수 있는 이중틀 방안을 검토하였다. 6차년도 사업체패널조사의 횡단면 모집단에 대한 표본추출틀은 2014년 12월 말 기준의 「전국사업체조사」의 사업체 명부 중 일부 업종을 제외한 상시근로자 30인 이상인 사업체이며, 조사수행, 자료분석 등을 고려해 일부 분류를 추가로 제외한 총 57,894개 사업체로 구성되었다. 모집단 층화는 산업분류와 사업체 규모 대상 40개 층으로 구성하였고, 비례배분과 등배분의 절충형태인 키쉬배분(Kish allocation)을 적용하였다.

기존 사업체에 대한 종단면 패널가중치는 1차년도 조사의 횡단면 가중치를 기초가중치로 고려하여 이후 각각의 추적조사에서 응답성향에 대한 모형적합을 통해 패널이탈과 무응답을 보정해 산출하였다. 6차 추적조사의 패널가중치 산출 역시 기존 패널가중치 산출방법과 동일한 방법이 이용되었으며, 가중치 산출 대상은 1차에서 5차 조사까지 모두 응답한 총 898개 사업체들이다

기존표본과 신규표본의 횡단면가중치 산정 역시 기존방식과 동일

하게 적용하였다. 최초년 혹은 2015년 표본추출 시 사용하였던 산업 대분류와 사업체 규모를 층으로 하는 복합표본설계(complex sample design)를 반영한 설계가중치, 무응답 조정, 사후층화조정 등의 세 가지 구성요소를 통합하여 산정하였다. 기존표본과 신규표본의 횡단면 가중치는 의사최대가능도 추정(Skinner and Rao, 1996)에 근거하여 산출되었다.

### 3. 웹을 활용한 조사 효과에 대한 실증연구 II

한국노동연구원 노동패널팀은 2017년, 표본이탈(attrition)에 따른 대표성의 한계극복을 위해 현재 진행되고 있는 CAPI(Computer Assisted Personal Interviewing)의 대안으로 웹을 기반으로 한 CAWI(Computer Assisted Web Interviewing) 도입을 검토한 바 있다. 1차 연구에서 조사도구별, 즉 CAWI와 CAPI 조사방법에 따른 항목별 결과의 차이를 검증했고, 2차 연구에서는 조사도구별 유지율의 차이와 표본이탈의 편향발생 여부, 1·2차 웹조사 간 실험 결과 차이, 수정된 설문문항에 따른 결과 차이 분석을 대상으로 연구를 진행하였다.

표본가구의 유지율에 있어서는 CAWI의 유지율이 CAPI에 비해 다소 높은 경향을 보이며, 개인은 유사한 경향을 보이나 가구와 개인 모두 유지율에 있어서는 조사도구별 차이가 통계적으로 유의하지 않은 것으로 나타난다.

오류유형에 있어서는 ‘가구원 내 가구원번호의 오류’, ‘가구원 추가 시 가구원번호 오류’, ‘이름 오류’, ‘가구 전체 모두 다른 이름’ 등의 유형으로 구분하여 살펴본 결과, CAPI와 CAWI가 통계적으로 차이가 있다고 분석되었다.

조사도구별 패널 유지집단과 탈락집단 간 차이에 있어서는 탈락집단의 가구원 수, 지역분포, 성별, 학력에서 통계적으로 유의한 차이가 나타나지 않은 반면, 연령별로는 CAWI 조사대상의 탈락자 평균 연령이 CAPI에 비해 통계적으로 유의하게 상대적으로 낮은 경향을

보이고 있다.

가구 및 개인의 특성 차에 있어 가구소득의 경우, 1차년도에는 금융소득액을 제외하고 CAPI의 응답률이 높게 나타났으며, 근로소득액, 부동산소득액 및 이전소득액의 응답률은 조사도구 간 통계적으로 유의한 차이가 있는 것으로 분석되었다. 또, 2차년도 조사에 있어서는 기타소득액을 제외하고 모든 항목에 있어 CAPI의 응답률이 높게 나타나고 있다. 통계적 유의성이 나타나는 항목은 1차조사와 동일하게 부동산소득액, 이전소득액과 소득총액이다. 그 외 지난 한 달 소득, 월평균 저축액에 있어 CAPI의 응답률이 통계적으로 유의하게 높은 것으로 나타났으며, 생활비 및 주거 관련 변수, 가구 생활비, 일 자리 관련 변수, 사회보험 수급, 혼인상태 변화 여부 등에서는 유의한 차이가 나타나지 않는다. 다만, 흡연 및 음주 응답비중에 있어서는 CAWI가 통계적으로 유의하게 CAPI의 응답률보다 높은 것이 특징이다.

직장만족도, 현재 건강상태에 있어서는 1차, 2차년도 조사 모두 CAWI 집단의 평균점수가 CAPI 집단에 비해 통계적으로 유의하게 낮은 것으로 분석되었다.

2차년도 조사 시 1차 조사 대비 기본 설문항목의 배열 화면을 변경하여 육아휴직급여, 출산전후휴가급여의 응답률에 차이가 있는지 확인해 본 결과를 살펴보면, 1차년도에는 휴업급여, 육아휴직급여, 출산전후휴가급여, 기타 항목에서 CAPI 집단보다 CAWI 집단의 응답률이 통계적으로 유의하게 높은 것으로 나타났다. 반면 2차년도에는 노령연금, 육아휴직급여, 출산전후휴가급여, 실업급여, 기타 항목에서 CAWI 집단의 응답률이 통계적으로 유의하게 높은 것으로 나타난다.



#### 4. 출산력변수 형성방법 및 문제점 진단

제4장에는 주요 인구통계학적 변수로 꼽히는 출산 관련 변수에 주목하여 출산력 구축 방안을 모색하는 한편, 출산력 구축 시 발생하는 문제점을 파악하기 위한 연구결과가 담겨 있다.

출산력변수는 관련변수들을 다수 결합해 추가 출생 자녀를 파악하여야 하는 만큼, 변수별 오류들이 집적되어 연구자의 자의가 반영될 가능성이 크다. 본 연구에서는 출산력 생성을 위해 사교육 문항을 이용하는 한편, 가구주와의 관계 변수와 결합하여 추가 출산이력을 추적하고, 신규조사자의 경우 응답한 출산이력 정보를 업데이트하는 방식으로 총 자녀수 변수를 생성하였다. 더불어 사교육 문항 대신 신규 가구원의 편입 사유 문항을 사용하는 방안 역시 고려하였으나 여전히 한계점은 노출된다. 출산력 구축을 위한 개선 방안에 있어서는 매년 출산이력의 변화를 직접 질문하는 것이 효과적 대안으로 보인다.

#### 5. 사업체패널조사의 경영의사결정 참여 설문문항들의 효과적 활용 탐색

본 연구는 근로자 대표조직의 경영 참여를 측정하는 총 11개 설문 문항들의 신뢰도와 타당성을 검토한다. 경영 참여 설문문항들이 측정하고자 하는 근로자 대표조직의 경영 참여가 가지는 이론적·현실적 중요성에도 불구하고 사업체패널조사 내 노사관계 영역들의 다른 설문문항들에 비해 그 활용도가 높지 않다.

본 연구는 광의와 협의의 경영 참여로 정의하고, 사업체의 경영 참여 수준을 측정하는 본 연구의 실증분석은 크게 세 부분으로 구분할 수 있다. 첫째, 경영의사결정 참여 설문문항들에 대한 신뢰도를 검증한다. 둘째, 경영 참여 지수의 타당도를 검증하기 위하여 경영 참여 수준, 선행변수들, 성과변수들 간 상관관계를 분석한다. 참여지수는 총 11개 영역에서 참여하는 데 동일한 가중치를 부여하는 단순합과

영역별 참여 난이도에 따른 가중치를 부여하는 Rasch 추정값으로 측정하였다. 선행변수들로 노사 양측이 측정한 노사의 노사관계정책(예, 회사의 파트너십과 노동조합의 협조전략), 노사관계 결과변수로는 노사관계분위기와 고용안정성을, 성과변수로는 정성적·정량적으로 측정된 노동생산성과 재무성과를 포함하였다. 마지막으로, 패널 고정효과 모형을 이용하여 경영 참여 수준과 선행변수들, 그리고 성과변수들과의 관계를 분석한다. 결론에서는 본 연구의 분석결과들과 함의를 논의한 후 향후 경영의사결정 참여 설문문항을 분석에 사용하기 위한 제안 사항들을 논의한다.

측정과 관련하여 가장 중요한 지표 중의 하나로서 신뢰도(reliability)는 측정 또는 관찰 값들이 동일하거나 일관된 결과를 보이는 정도를 나타내는 개념이다. 본 연구는 사업체패널조사의 경영의사결정 설문문항들의 신뢰도를 추정하기 위하여 명목척도를 분석의 도구로 활용한다. 2013년 노사 간 일치도와 가정을 완화하여 Spearman's rho를 통해 노사 양측 응답의 일관성을 신뢰도로써 평가한다. 또한 총 11개의 경영의사결정 참여 설문문항을 사업체의 경영 참여라는 잠재 특성을 반영하는 단일 요인으로 가정하고, 참여지수를 구성하기 위한 일관성을 가지는지 내적 일관도를 통해 평가한다.

타당도는 측정도구가 측정하고자 하는 대상을 충실히 측정한 정도로 정의할 수 있다. 경영 참여 변수가 이론적으로 관계가 있는 변수와의 상관관계를 측정하는 수렴(convergent) 타당도에 초점을 두며, 이 경우 상관관계가 높을수록 타당도가 높다고 할 수 있다. 본 연구는 근로자 대표조직의 경영 참여와 이론적으로 관계가 있는 선행변수들과 성과변수들과의 상관관계를 평가하고, 패널고정효과 모형을 통해 다른 변수들의 영향을 통제한 후에도 경영 참여 수준과 선행 및 성과변수들의 관계가 일반화될 수 있는지를 검토한다.

본 연구는 2005~2015 사업체패널조사에서 사업체 내 근로자의 대표조직, 즉 노동조합 또는 노사협의회가 조직되어 있으며 3회 이상 연속으로 응답한 사업체들을 선정하였다. 따라서 총 12,353개 사업체

중에서 경영의사결정 참여수준을 기준으로 7,192개의 사업체들이 기본적인 분석을 위한 표본으로 선정되었으며, 통제변수와 종속변수들의 결측치에 따라 다른 수의 표본이 분석에 이용되었다.

2005년부터 2015년까지 조사된 총 11개 영역에서 근로자 대표조직의 경영의사결정 참여 여부를 1) “노사협의”와 “공동결정”, 그리고 2) “공동결정”과 같이 광의와 협의의 차원으로 정의하였다.

경영의사결정 참여수준의 선행변수로 사측과 노측 각각의 노사관계정책을 측정하였다. 성과변수로는 정성적 지표로서 노사관계분위기, 고용안정성, 노동생산성과 재무적 성과를 측정하였다. 노사관계분위기는 단일문항과 노사 양측의 복합문항을 이용하여 측정하였으며, 고용안정성은 해고가 없으면서 합의사항이 있는 경우로 측정하였다. 정성적 성과지표로는 전년 기준 동종업체 평균과 비교한 노동생산성과 재무적 성과를 사용하였으며, 정량적 지표로는 1인당 부가가치와 1인당 영업이익을 측정하였다.

경영의사결정 참여 설문문항들의 신뢰도 검증결과는 다음과 같다. 2013년 노사별 평가자 간 일치도가 11개 영역에서 53~63%로 나타났다. 서열척도로 추정한 Spearman's rho는 0.35~0.65로 나타났다. 2005년부터 2015년까지 사측 응답의 내적 일관도는 협의의 참여가 연도별로 0.83부터 0.92, 광의의 참여가 0.79부터 0.91로 나타나 경영 참여 수준을 측정할 수 있는 지수를 사용하기에 적합한 것으로 나타났다.

두 가지 방식으로 정의된 경영 참여를 기반으로 사업체 내 근로자 대표조직의 경영 참여 수준을 측정하기 위하여 단순합과 Rasch 추정값을 참여지수로 사용하였다. 먼저 참여지수 간 상관관계를 살펴보면, 동일한 참여지수 내 상관관계는 약 0.98, 참여지수 간 상관관계는 약 0.35~0.36으로 나타났다. 경영 참여 수준과 노사관계정책들 간 상관관계는 통계적으로 유의하였으며, 특히 사측 응답으로 계산된 참여지수와 노측 응답으로 계산된 이 노사관계정책들 간 상관관계가 더 높게 나타났다. 경영 참여 수준과 성과지표들 간 상관관계를

살펴보면, 참여지수가 정성적 성과지표보다 정량적 성과지표가 높은 상관관계를 보여주었다.

마지막으로, 근로자 대표조직의 경영 참여 수준과 선행변수들, 그리고 성과변수들과의 관계를 패널고정모형을 이용하여 분석한 결과는 다음과 같다. 노동조합의 조직 여부가 경영 참여 수준의 중요한 결정요인으로 나타났다. 다음으로 참여지수 1의 결정요인으로 회사의 파트너십과 노동조합의 협조전략이 모두 통계적으로 유의하였으나, 참여지수 2의 결정요인으로는 협조전략만이 통계적으로 유의하였다.

다음으로, 사업체 성과의 결정요인으로 경영 참여 수준의 영향을 요약하면 다음과 같다. 노사관계분위기와 관련하여 참여지수 2가 참여지수 1보다 통계적으로 유의한 영향을 미치는 것으로 나타났으며, 이는 노사 양측 모두에게서 일관되게 나타났다. 한편 참여지수 1이 높아질수록 고용안정성 또한 통계적으로 유의한 수준에서 높아지는 것으로 나타났으나 참여지수 2는 그렇지 않았다.

정성적 성과지표와 관련하여 참여지수 1은 노동생산성을 통계적으로 유의한 수준에서 높이는 것으로 나타났으나 재무성과에서는 그렇지 않았다. 참여지수 2는 두 정성적 성과지표들과 통계적으로 유의하게 나타났다. 반면 참여지수 1과 참여지수 2 모두 정량적 성과지표와 통계적으로 유의하지 않은 것으로 나타났다. 이는 상관관계 분석결과와는 다르지만 사업체패널조사를 분석한 국내 연구들의 결과와는 유사한 것으로 볼 수 있다. 위 결과를 요약하자면, 협의의 경영 참여가 광의의 경영 참여보다, 그리고 Rasch 참여지수가 단순합 참여지수보다 결정요인 분석에서 일관되고 통계적으로 유의한 것으로 나타났다.

후속 연구가 경영 참여의 하위요인 구조를 뒷받침할 이론적 논의와 실증분석을 진행할 필요가 있다. 본 연구에서는 하위요인 구조 분석을 진행하지 않았는데, 그 이유는 단일요인으로서 경영 참여가 갖는 신뢰도와 타당도를 검증하는 것이 목적이었기 때문이다. 또 단일요인으로서 참여지수를 구성하는 데 경영 참여 측정이 충분한 신뢰

도를 가진 것으로서 검증되었기 때문이다. 가장 유망한 경로는 전략적 선택이론을 기반으로 사업체 내 경영의사결정 참여를 전략적 수준과 기능적 수준으로 구분하고, 하위요인 구조가 수준별 참여들을 반영하는가를 검증하는 것이다.

다음으로, 사업체패널조사의 정보를 최대한 활용하는 분석기법들, 예를 들어 동적패널분석(dynamic panel modeling)과 잠재성장모형(latent growth modeling) 등을 활용할 수 있는 연구 질문들을 발전시키고 실증할 필요가 있다. 무노조사업체에서 노사협의회의 경영 참여가 드러내는 성장패턴과 그 결정요인 파악을 위한 연구를 예로 들 수 있겠다.

마지막으로, 경영의사결정 참여 설문문항과 관련된 한계를 언급할 필요가 있다. 먼저 유노조사업장 내 노사협의회의 경영 참여가 측정되지 않고 있다. 이로 인해 근로자 대표조직인 노동조합과 노사협의회의 경영 참여보다 노동조합 유무에 따른 경영 참여와 분석적으로 구분하기 어려운 한계가 존재한다.

또한 노동조합의 경영 참여와 관련하여, 응답자가 설문문항들을 어떻게 해석하고 응답하는지 파악할 필요가 있다. 특히 기능적 수준에서, 예를 들어, 임금제도 변경과 관련한 노사합의에 대하여 응답자가 “노사협의”로 응답할지 “공동결정”으로 응답할지 모호한 점이 있다. 사업체패널조사에서는 상당수의 노동조합 표본에서 “노사협의”로 응답이 이루어졌다는 점에서 그 의미를 명확히 할 필요가 있다.



## 제 1 장 서 론

본서는 「패널자료 품질개선 연구 시리즈」의 8차 연구 결과물이다.

한국노동연구원 패널데이터연구실은 생산·품질제고·연구분석 등 전 과정에 걸쳐 전담 관리하는 한국노동패널조사(Korean Labor and Income Panel Studies : KLIPS)와 사업체패널조사(Workplace Panel Study : WPS)의 품질개선을 위해 2012년도에 시작해 현재까지 8차에 걸쳐 연구를 수행하고 있다. 패널자료 품질개선 연구의 주제는 표본의 설계·추출, 가중치 조정, 표본이탈·무응답과 히핑 및 측정오차의 보정, 자료검증과 기존 항목을 이용한 변수추정 등 패널조사자료의 품질개선을 위한 다양한 세부 주제들로 구성되어 있다. 패널자료 품질개선 연구의 지난 차수별 세부 연구주제는 <표 1-1>과 같다.

「패널자료 품질개선 연구 VIII」은 지난 패널자료 품질개선 연구 시리즈의 연장선에서 ‘제6차 사업체패널조사 가중치 산출 연구’, ‘웹을 활용한 조사 효과에 대한 실증연구 II’, ‘출산력변수 형성방법 및 문제점 진단’, ‘사업체패널조사의 경영의사결정 참여 설문문항들의 효과적 활용 탐색’이라는 주제를 다루고 있다.

〈표 1-1〉 패널자료 품질개선 연구(I~VII) 세부 연구주제

**패널자료 품질개선 연구 (I)**

- 사업체패널자료 가중치조정 연구

**패널자료 품질개선 연구 (II)**

- 개인조사자료와 사업체조사자료 간 임금분포의 히핑현상 비교
- 조사방법과 히핑의 관계
- 노동패널 표본이탈 연구
- 사업체패널조사에 나타난 표본이탈 현황 및 대체표본에 대한 분석
- 고령화연구패널조사 건강영역의 품질개선 방안

**패널자료 품질개선 연구 (III)**

- 소득 히핑의 보정
- 측정오차에 대한 통계적 보정방법론 연구
- 표본이탈 교정을 위한 계량경제학 방법
- Nonignorable panel attrition에 대한 보정 연구
- Paradata를 활용한 패널 무응답 보정을 위한 가중치 부여 방안

**패널자료 품질개선 연구 (IV)**

- 동태적 패널모형에서 표본이탈 교정
- Paradata를 이용한 무응답 자료 회귀분석
- 사업체패널조사의 표본 설계 관련 연구
- 한국에서의 자료검증 연구의 필요성에 대하여

**패널자료 품질개선 연구 (V)**

- 사업체패널조사의 시점 및 표본 대표성
- 사업체패널조사의 표본이탈과 대표성
- 표본 이탈률의 크기와 편향의 관계

**패널자료 품질개선 연구 (VI)**

- 사업체패널조사의 시점 및 표본대표성
- 사업체패널조사의 표본이탈과 대표성
- Propensity 점수를 이용한 사업체조사의 대표성 연구
- 표본 이탈률의 크기와 편향의 관계
- 2016 사업체패널조사의 추가표본설계 및 표본추출
- 노동패널 항목 무응답 처리방법 보완
- 사업체패널조사의 신뢰도와 타당도 검토
- 노동패널자료 소득세액 추정



〈표 1-1〉의 계속

**패널자료 품질개선 연구 (Ⅶ)**

- 한국노동패널조사 표본 추가 연구
- KLIPS 보육 관련 설문조사의 문제와 데이터 품질 제고 방안
- 소득세 추정 모형의 필요성과 활용 방안
- 한국노동패널조사와 경제활동인구조사에 나타난 노동시장 활동 비교
- 한국노동패널조사 웹서베이 도입 가능성 검토를 위한 연구
- 웹을 활용한 조사의 효과에 대한 실증연구: 노동패널 CAWI 실험 데이터의 분석

사업체패널조사(WPS : Workplace Panel Survey)는 2006년 패널표본이 개편된 이후 2014년까지 격년 주기로 총 5차례의 조사가 진행되었다. 그러나 조사가 진행됨에 따라 휴폐업, 응답 거절 등으로 인한 표본마모(attrition)가 지속적으로 이루어지면서, 대표성 확보를 위한 표본대체가 지속적으로 이루어진 바 있다. 하지만 이 역시 누적된 표본대체와 사업체의 특성변화로 인해 사업체패널조사의 지속가능성과 동태적 변화 파악에 있어 한계에 봉착하였으며, 이에 2016년 6차 조사에서 신규 표본사업체를 대규모 추가하였다. 결과적으로, 사업체패널조사는 기존 패널표본에 최신 모집단에서 추출된 추가표본이 혼재한 분리패널(split panel)의 형태를 띠게 되었으며, 이에 변화된 표본구성을 반영하고자 적절한 표본 가중치의 산출을 위한 연구 필요성이 제기되었다. 이제 제2장에서는 제6차 사업체패널조사를 위한 종단면가중치에 대해 논의하고, 개별적 횡단면가중치와 이중추출틀 접근에 기초한 통합표본의 유사최대가능도 횡단면가중치의 산출 과정 및 결과에 대해 논의하도록 한다.

『패널자료 품질개선 연구 시리즈 Ⅶ』에서는 한국노동패널조사의 웹서베이 도입 가능성과 이의 활용성 및 효과에 대한 실증연구를 수행한 바 있다. 앞서 언급한 사업체패널조사의 예와 같이, 패널조사는 주기적으로 이루어지는 조사의 특성상 비성공 가구의 증가로 인한 패널 마모가 필연적으로 발생하게 되며, 한국노동패널조사 역시 예외는 아니다. 강력 거절 및 이사 추적 실패의 비중이 높아짐에 따라, 한국노동패널조사는 패널유지를 제고를 위한 대안의 일환으로 웹서베이 방식의 도입가능성을 모색해 왔다. ‘웹을 활용한 조사 효과에 대한 실증연구 Ⅱ’는 패널자료 품질개선 연구

(VII)의 웹조사 관련 연구에 이은 2년차 성과로서, 표본이탈에 CAPI-CAWI 조사그룹 간 이질성이 자료의 신뢰성 및 대표성에 미치는 영향을 분석하였다. 즉, 본 연구의 목적은 1차 실험조사와 2차 실험 조사의 결과를 비교 분석하여 조사도구별 차이를 파악하고, CAWI가 대면조사인 CAPI의 한계를 극복할 수 있는 대안으로 가능한 것인가를 검토하는 데 있다.

한국노동패널조사는 본 설문에 연동하여 제공하는 raw data 외 직업력(work history) 데이터를 제공하고 있으며, 패널조사의 신뢰성 및 대표성 검토, 사용자 친화적 데이터 구축을 위한 노력의 일환으로 주거력, 결혼력 등의 구축과 관련한 연구를 수행해 온 바 있다. 본 「패널자료 품질개선 연구 시리즈 VIII」은 주요 인구통계학적 변수로 꼽히는 출산 관련 변수에 주목하고, 출산력 구축 방안 및 이에 대응하여 발생하는 문제점을 파악하기 위한 연구를 수행하였다. 저출산 문제가 심화되고 있는 최근, 그 중요성이 더욱 높아지고 있는 출산 관련 변수는 자녀의 출생이 여성근로자들의 경력단절 및 임금에 미치는 영향, 일·가정 양립정책과 출산 영향 등 주요 연구에 있어 필수적 요소로 자리 잡고 있다. 이에 본 연구는 출산력 자료 구축 시 활용될 수 있는 변수들과 구축방법들을 정리하여 연구자의 출산력변수 이해를 돕는 한편, 출산력변수 생성 과정에서 발생하는 여러 문제점들을 토대로 노동패널 출산력 데이터의 품질제고를 위한 개선 방안을 제시하고자 한다.

제5장은 사업체패널조사의 경영의사결정 참여와 관련한 설문문항의 신뢰도 및 타당성을 검토한다. 경영 참여 설문문항들이 측정하고자 하는 근로자 대표조직의 경영 참여가 갖는 이론적·현실적 중요성에도 불구하고 사업체패널조사 내 노사관계 영역들은 다른 설문문항들에 비해 그 활용도가 높지 않았다. 본 연구의 성과가 관련 문항의 활용도를 높이는 데 기여할 것으로 기대되는 이유이다. 본 연구의 내용은 경영의사결정 참여 설문문항들에 대한 신뢰도를 검증함으로써 경영 참여 지수의 타당도 검증을 위한 경영 참여 수준·선행변수들·성과변수들 간 상관관계 분석, 패널고정효과 모형을 이용한 경영 참여 수준과 선행변수들 및 성과변수들과의 관계 분석으로 나뉜다.

『패널자료 품질개선 연구 시리즈 I~Ⅷ』은 한국노동패널조사가 국내 패널자료의 선구적 역할 및 대표 조사자료로서의 위상을 제고하고 중장기적 관점에서의 발전방안을 모색하는 한편, 종단면 조사 자료로서의 활용성 및 품질개선을 위한 노력의 일환으로 수행되어 왔다. 한국노동연구원 한국노동패널팀은 한국노동패널조사가 한국을 대표하는 패널조사로서 자리매김하고, 한국의 경제 및 노동시장에 관련한 다양한 논의의 개발과 정책이슈의 개발 및 수립, 관련 연구의 지평 확대에 노력을 경주할 것을 약속드린다.

## 제 2 장

## 제6차 사업체패널조사 가중치 산출 연구

## 제1절 서 론

사업체패널조사(WPS : Workplace Panel Survey)는 한국노동연구원이 주관하여 수행하는 조사로 우리나라 사업체를 대표하는 패널을 구성하여 격년주기로 추적조사하는 종단면조사(longitudinal survey)이다. 우리나라 고용구조 및 노동수요는 물론 인적자원관리체계를 과학적이고 체계적으로 평가하여 바람직한 노동시장정책을 수립하는 기초자료를 제공하는 것이 조사의 목적이다.

본 조사는 2006년 패널표본이 개편된 이후 2014년까지 격년 주기로 총 5차례의 조사가 진행되었다. 최초 조사 이후 4차례에 걸친 추적조사 과정에서 휴폐업, 응답 거절 등으로 인한 표본패널사업체의 탈락(attrition)이 계속적으로 진행되면서 대표성 확보를 위해 패널이탈사업체들에 대한 지속적인 표본대체가 실시되었다. 하지만 누적된 표본대체와 사업체 특성 변화로 인해 사업체패널조사의 지속가능성은 물론 동태적 변화를 파악함에 있어서의 적절성 문제가 꾸준히 제기되어 왔다. 이에 따라 2016년에 수행한 6차 조사에서는 신규 표본사업체가 추가된다(박인호·김기민, 2016).

2016년 조사결과는 기존 패널모집단의 특성을 포함해 표본 추가시점을 기준으로 한 횡단면모집단의 특성을 모두 반영할 수 있어야 한다. 따라서

6차 조사를 위한 표본가중치 산출에서는 다음의 두 가지 측면이 반드시 고려되어야 한다.

- (i) 기존 패널모집단의 대표성과 신규모집단 대표성
- (ii) 두 시점 간 사업체 생멸 및 표본 중복의 포함을 반영한 추출확률

제6차 사업체패널조사는 기존 패널표본에 최신 모집단에서 추출된 추가표본이 더해지는 복잡한 형태로 구성되었다. 즉, 2006년 이후 지속된 기존패널과 2016년부터 시작되는 추가패널이 혼재한 분리패널(split panel) 방식이다. 이러한 표본구성을 반영하기 위해 이중추출틀(dual frame) 접근하에 의사최대가능도 이론(pseudo-maximum likelihood theory)을 이용한 표본가중치를 개발하였다. 이중추출틀 접근은 기존 패널과 추가표본이 각각 대표하는 모집단을 (간접적으로) 분리하여 평가할 수 있어 생성(birth) 및 폐업(death) 등의 시점 간 모집단 변동도 확인할 수 있게 해준다. 본 연구에서는 이중추출틀 구조하에서 일반적으로 표본가중치가 반영해야 할 추출확률, 무응답조정, 이상치 가중값 조정, 모집단 보조정보의 반영 등이 복합적으로 반영될 수 있는 적절한 가중치 산출방법을 적용하고자 한다. 제2절에서는 기존 표본설계 및 응답현황에 대해 기술한다. 제3절에서는 6차 조사에서 고려한 추가표본을 위한 표본설계 내역을 살펴보고 조사응답현황에 대해 기술한다. 제4절에서는 6차 사업체패널조사를 위한 종단면가중치에 대해 논의하고, 개별적 횡단면가중치는 물론 이중추출틀 접근에 기초한 통합표본의 유사최대가능도 횡단면가중치의 산출 과정 및 결과에 대해 논의한다.

## 제2절 기존 표본설계 및 응답현황

## 1. 기존 표본설계

## 가. 1차년도 표본설계

사업체패널조사는 이전 3차례의 선행적 시행을 바탕으로 2006년에 새롭게 1차년도 표본설계가 수행되었다. 조사단위는 사업장<sup>1)</sup>이며, 농림어업 및 광업 등의 일부 산업을 제외한 전 산업에서 상시근로자 30인 이상 규모의 모든 사업체를 대상으로 한다. 통계청의 「전국사업체조사(舊 사업체기초통계조사)」자료를 표본추출틀로 하여 민간부문 3,916개 사업체를 층화추출에 의해 선택하였다. 공공부문에 대해서는 약 365개의 별도 표본을 선정하였다.

<표 2-1>은 공공부문을 제외한 1차 조사에 대한 산업분류와 함께 사업체 규모에 따른 표본수와 추출률을 보여주고 있다. 표본설계 당시 근로자 기준 500인 이상의 사업체에 대해서는 전수추출을 하였으나, 실제 조사를 진행하는 과정에서 산업·규모·지역별 목표응답수가 채워져 해당 셀에 대한 조사가 종료되면 셀에 남아있는 사업체는 표본에서 제외하였다.

<표 2-2>는 1차년도 사업체조사의 공공부문과 민간부문의 조사현황이다. 먼저, 공공부문 표본은 기획재정부(舊 기획예산처)가 선정한 공공기관 308개<sup>2)</sup>와 51개 지방공기업인 총 359개를 대상으로 조사하였으며, 이때 특별조사 형식으로 1차년도 조사에서만 전체 기관을 대상으로 조사를 실시하였다.<sup>3)</sup> 조사부적격인 31개를 제외한 유효표본 328개 기관 중

- 
- 1) 공공부문, 금융·보험업, 전기·가스·수도업은 기업(혹은 기관) 단위를 기준으로 조사하였다.
  - 2) 2006년에 314개 기관이 기획재정부에 의해 공공기관으로 지정되었는데, 사업체패널조사에는 이 중 6개 기관에 대해서 예비조사를 실시함에 따라 본조사 표본에서 제외하였다.
  - 3) 2차 조사부터는 이 중 일부 기관에 대해서만 조사를 실시하였다.

290개 기관이 응답하였다. 민간부문의 표본은 총 3,916개소로 이 중 692개소의 사업체가 조사부적격인 것으로 파악되었고, 유효표본 3,224개 중 1,615개소의 사업체가 응답하였다.

〈표 2-1〉 공공부문을 제외한 사업체패널 1차 조사의 표본수 및 추출률

(단위: 개소, %)

		30~99인	100~299인	300~499인	500인 이상	전 체
제조업	경공업	207 ( 5.4)	119 ( 15.3)	95 (84.8)	59 (95.2)	480 (10.0)
	화학공업	140 ( 5.7)	98 ( 18.0)	74 (88.1)	56 (86.2)	368 (11.7)
	금속·자동차·운송	214 ( 4.6)	116 ( 13.1)	75 (74.3)	99 (85.3)	504 ( 8.7)
	전기·전자·정밀	147 ( 6.3)	100 ( 15.5)	80 (89.9)	98 (85.2)	425 (13.3)
건설업		170 (11.7)	97 ( 48.3)	13 (68.4)	15 (71.4)	295 (17.4)
개인서비스업		181 ( 6.5)	94 ( 18.5)	57 (96.6)	41 (91.1)	373 (10.9)
운수업		106 ( 5.4)	85 ( 8.3)	49 (61.3)	33 (71.7)	273 ( 8.7)
통신업		35 (39.3)	23 (100.0)	9 (75.0)	10 (100.0)	77 (57.5)
금융보험업		57 (29.7)	23 ( 45.1)	19 (55.9)	26 (59.1)	125 (38.9)
사업서비스업 기타		200 ( 5.8)	102 ( 11.5)	93 (53.8)	111 (82.8)	506 (10.9)
사회서비스업		182 ( 4.1)	94 ( 12.8)	79 (81.4)	112 (73.7)	467 ( 8.6)
전기·가스·수도업		13 (48.1)	10 ( 83.3)	-	-	23 (59.0)
전 체		1,652 ( 6.0)	961 ( 15.3)	643 (74.8)	660 (81.5)	3,916 (11.0)

주: ( ) 안은 추출률.

자료: 김수진·김기민(2007), 『노동리뷰』 2007. 2(통권 제26호), 한국노동연구원.

〈표 2-2〉 2006년 사업체패널조사 조사 현황

(단위: 개소, %)

			전 체		민간부문		공공부문	
			수	비 율	수	비 율	수	비 율
조 사 부 적 격	30인 미만 사업체		393	9.2	368	9.4	25	7.0
	해당 안 됨(2006년 설립 등)		23	0.5	20	0.5	3	0.8
	폐업/폐쇄/휴업/합병		63	1.5	62	1.6	1	0.3
	결번		64	1.5	64	1.6		
	잘못된 번호		98	2.3	97	2.5	1	0.3
	지방 공기업		30	0.7	30	0.8		
	리스트 중복		14	0.3	14	0.4		
	사업체/사업체명 다름		31	0.7	30	0.8	1	0.3
	기타 사항		7	0.2	7	0.2		
	소 계		723	16.9	692	17.7	31	8.6
조 사 실 패	거절		1,283	30.0	1,245	31.8	38	10.6
	연락 안 됨	재연락	50	1.2	50	1.3		
		담당자 부재	45	1.1	45	1.1		
	기타 사항	내부결제 중	9	0.2	9	0.2		
		보류	192	4.5	192	4.9		
		진행중단	68	1.6	68	1.7		
	소 계		1,647	38.5	1,609	41.1	38	10.6
조사 성공		1,905	44.6	1,615	41.2	290	80.8	
전 체		4,275	100.0	3,916	100.0	359	100.0	

자료: 김수진·김기민(2007), 『노동리뷰』 2007. 2(통권 제26호), 한국노동연구원.

## 나. 2~5차년도 표본추가

1차년도 조사 이후 격년 주기의 2~5차년도 추적조사를 진행하면서 패널사업체의 휴·폐업, 응답 거절 등의 사유로 인한 표본탈락이 진행되었다. 이에 각 조사시점별로 횡단면적 대표성을 유지하기 위한 목적에 따라 표본추가 작업이 진행되었다. 표본추가는 원패널 구축 당시와 동일하게 조사대상, 목표모집단, 조사가능모집단, 기본단위, 추출단위 등을 고려하여 수행하였다(이기재, 2015).



## 2. 기존 조사 응답현황

### 가. 1차년도 조사

<표 2-3>은 2006년 사업체패널조사의 부적격률<sup>4)</sup>과 응답률<sup>5)</sup>을 사업체

<표 2-3> 2006년 사업체패널조사(1차 조사)에 대한 표본 사업체의 응답률 및  
표본 부적격률

(단위: 개소, %)

		응답 (A)	조사 실패 (B)	소계 (C=A+B)	응답률 (A/C)	표본 부적격 (D)	표본 부적격률 (D/T)	합계 (T)	
공공부문		290	38	290	88.4	31	8.6	359	
규모	30~99인	622	586	652	51.5	444	26.9	1,652	
	100~299인	408	450	508	47.6	103	10.7	961	
	300~499인	324	248	624	56.6	71	11.0	643	
	500인 이상	261	325	761	44.5	74	11.2	660	
산업	제조업	경공업	185	195	185	48.7	100	20.8	480
		화학공업	151	173	151	46.6	44	12.0	368
		금속·자동차·운송	209	226	209	48.0	69	13.7	504
		전기·전자·정밀	161	191	161	45.7	73	17.2	425
	비제조업	건설업	107	112	107	48.9	76	25.8	295
		전기·가스·수도업	20	3	20	87.0	0	0	23
		개인서비스업	152	152	152	50.0	69	18.5	373
		유통서비스업	172	107	172	61.7	71	20.3	350
		사업서비스업	243	251	243	49.2	137	21.7	631
		사회서비스업	215	199	215	51.9	53	11.4	467
전 체		1,876	1,934	2,376	53.6	766	172.6	4,576	

주: 공공부문은 표본추출 시 규모와 산업에 대한 정보가 없음.

자료: 김수진·김기민(2007), 『노동리뷰』 2007. 2(통권 제26호), 한국노동연구원.

- 4) 부적합률은 조사표본 전체에서 조사대상으로 부적합하여 조사대상에서 제외한 표본의 비율을 말한다.
- 5) 응답률은 조사표본 전체에서 조사대상으로 부적합한 표본을 제외한 나머지 유효표본에 대한 조사 성공 표본의 비율을 말한다.

규모 및 산업분류별로 보여준다. 사업체 규모별로 보면, 30~99인 사업체의 경우 조사 시점 당시 사업체의 규모가 축소되어 30인 미만으로 분류되거나, 폐·휴업 등의 사유가 발생하여 조사대상에서 제외된 비율이 높았다. 공공부문은 사업체 특성 등에 있어 변화가 적고 전화번호 등의 정확한 연락 정보 수집이 용이하므로 부적격률이 8.6%로 낮았고, 응답률도 88.4%로 매우 높았다. 반면 민간부문을 산업별로 보면, 부적격률은 건설업이 25.8%로 가장 높고 전기·가스·수도업이 0%로 가장 낮았다. 사업체 규모나 산업분류에 따른 응답률이 크게 저조한 부분은 나타나지 않았다.

간단히 정리하면, 1차 조사에서 민간부문 3,916개 사업체와 공공부문 359개 기관을 대상으로 조사를 실시하였고, 이 중 민간부문 1,615개소 사업체, 공공부문 134개 기관<sup>6)</sup>이 조사에 응답하여 총 1,749개의 기준년 패널사업체가 구축되었다.

#### 나. 2~5차년도 추적조사 및 표본추가

이렇게 구축된 패널사업체에 대해 2008년부터 격년마다 추적조사를 실시하였으며, 추수가 진행됨에 따라 휴·폐업 및 추적 불가로 인한 소멸사업체<sup>7)</sup>와 응답 거절 등의 사유로 인한 탈락사업체를 보완하고자 대체 표본 설계를 병행하며 조사를 지속해 나갔다.

<표 2-4>는 패널사업체를 구축한 후 2~5차에 걸친 추적조사에서 나타난 사업체 응답현황과 패널유지율을 보여주고 있다. 원표본 패널사업체(1,749개소)들에 대한 추적조사의 응답현황은 2차 조사의 1,415개소, 3차 조사의 1,229개소, 4차 조사의 1,091개소, 5차 조사의 960개소이다. (누적) 소멸사업체는 2차 조사의 115개소, 3차의 148개소, 4차의 172개소, 5차의 192개소 등이다. 따라서 실제 패널유지율은 2차 86.6%, 3차 76.8%, 4차 69.2%, 5차 61.7%이다. 더불어 대체사업체 수는 2차에 320개소, 3차에

6) 2006년에 공공부문 특별조사를 실시하여 최종 290개 기관에 대해 조사를 완료하였으나, 패널사업체는 134개로 선정하였다.

7) 사업체패널조사는 패널사업체에 대해서는 규모가 30인 미만으로 변경되더라도 추적조사하였다.

161개소, 4차에 238개소, 5차에 233개소 등으로 상세한 내역은 <표 2-6>을 참조할 수 있다.

<표 2-5>는 1차부터 5차 조사까지 대체표본을 포함한 사업체 수와 전 차수 대비 패널유지율을 보여주고 있다. 1차 표본사업체(1,749개소) 중 2차에 조사된 사업체 수는 1,415개소, 2차 표본사업체(1,735개소) 중 3차에 조사된 사업체 수는 1,508개소, 3차 표본사업체(1,737개소) 중 4차에 조사된 사업체수는 1,488개소, 4차 표본사업체(1,770개소) 중 5차에 조사된 사업체수는 1,537개소이다. 전 차수 표본 대비 실 표본유지율은 2차에 86.6%, 3차에 88.9%, 4차에 88.0%, 5차에 89.1%로 나타나 전 차수 대비 평균 약 88%의 높은 유지율을 보여주고 있다.

<표 2-4> 사업체패널조사 1~5차 패널사업체 추적조사 응답현황 및 패널유지율  
(단위: 개소)

	1차	2차	3차	4차	5차
패널사업체 수	1,749	1,415	1,229	1,091	960
응답 거절		219	153	114	111
(누적)휴폐업 및 추적 불가		115	33	24	20
패널유지율(%)		80.9	70.3	62.4	54.9
실 패널유지율(%) <sup>1)</sup>		86.6	76.8	69.2	61.7

주: 실 패널유지율은 휴·폐업 및 추적 불가 사업체를 제외한 사업체 수를 기준으로 계산함.

<표 2-5> 사업체패널조사 1~5차 전 차수 대비 표본유지율  
(단위: 개소)

	1차	2차	3차	4차	5차
전 차수 대비 응답사업체 수	1,749	1,415	1,508	1,488	1,537
응답 거절		219	189	203	188
휴폐업 및 추적 불가		115	38	46	45
표본유지율(%)		80.9	86.9	85.7	86.8
실 표본유지율(%) <sup>1)</sup>		86.6	88.9	88.0	89.1

주: 실 패널유지율은 휴·폐업 및 추적 불가 사업체를 제외한 사업체 수를 기준으로 계산함.

〈표 2-6〉 사업체패널조사의 1~5차의 전체 사업체에 대한 응답유형 현황

(단위: 개소)

	1차	2차	3차	4차	5차	민간	공공	전 체
1	○	○	○	○	○	861( 33.7)	99( 68.8)	960( 35.5)
2	○	○	○	○	×	126( 4.9)	5( 3.5)	131( 4.9)
3	○	○	○	×	×	131( 5.1)	7( 4.9)	138( 5.1)
4	○	○	×	×	×	135( 5.3)	6( 4.2)	141( 5.2)
5	○	×	×	×	×	258( 10.1)	5( 3.5)	263( 9.7)
6	○	×	○	×	×	13( 0.5)	0( 0.0)	13( 0.5)
7	○	×	×	○	×	0( 0.0)	0( 0.0)	0( 0.0)
8	○	×	×	×	○	3( 0.1)	0( 0.0)	3( 0.1)
9	○	○	×	○	×	8( 0.3)	1( 0.7)	9( 0.3)
10	○	○	×	×	○	1( 0.0)	0( 0.0)	1( 0.0)
11	○	×	○	○	×	8( 0.3)	1( 0.7)	9( 0.3)
12	○	×	○	×	○	0( 0.0)	0( 0.0)	0( 0.0)
13	○	×	×	○	○	0( 0.0)	0( 0.0)	0( 0.0)
14	○	×	○	○	○	43( 1.7)	3( 2.1)	46( 1.7)
15	○	○	×	○	○	28( 1.1)	7( 4.9)	35( 1.3)
16	○	○	○	×	○	0( 0.0)	0( 0.0)	0( 0.0)
17		○(추가)	○	○	○	186( 7.3)	0( 0.0)	186( 6.9)
18		○(추가)	○	×	×	65( 2.5)	0( 0.0)	65( 2.4)
19		○(추가)	×	○	×	0( 0.0)	0( 0.0)	0( 0.0)
20		○(추가)	×	×	○	0( 0.0)	0( 0.0)	0( 0.0)
21		○(추가)	○	○	×	28( 1.1)	0( 0.0)	28( 1.0)
22		○(추가)	○	×	○	0( 0.0)	0( 0.0)	0( 0.0)
23		○(추가)	×	○	○	0( 0.0)	0( 0.0)	0( 0.0)
24		○(추가)	×	×	×	41( 1.6)	0( 0.0)	41( 1.5)
25			○(추가)	○	○	103( 4.0)	1( 0.7)	104( 3.9)
26			○(추가)	○	×	24( 0.9)	0( 0.0)	24( 0.9)
27			○(추가)	×	○	1( 0.0)	0( 0.0)	1( 0.0)
28			○(추가)	×	×	32( 1.3)	0( 0.0)	32( 1.2)
29				○(추가)	○	206( 8.1)	0( 0.0)	206( 7.6)
30				○(추가)	×	32( 1.3)	0( 0.0)	32( 1.2)
31					○(추가)	224( 8.8)	9( 6.3)	233( 8.6)
32	○(공공)	×	×	×	×		148	148
합계 1						2,557(100.0)	144(100.0)	2,701(100.0)
합계 2						2,557	292	2,849

주: 1) 32번 유형은 공공부문 특별조사로 인해 발생한 경우임.

2) '합계1'은 1차 조사의 공공부문 특별조사로 인하여 조사된 사업체 수를 제외한 값이고, '합계2'는 이를 모두 포함한 값임.

<표 2-6>은 1차에서 5차 조사까지 전체 표본사업체의 응답유형 현황을 보여준다. 유형 1번부터 유형 16번까지는 원표본 패널사업체의 응답유형별 현황이다. 이전 차수에서 조사된 사업체 중 현 차수에 조사되지 않은 경우에도 이후 차수에서 조사를 수행하였음을 나타낸다. 1차 조사부터 5차 조사까지 최종 조사된 총 사업체의 수는 2,849개소이고, 이 중 민간부문은 2,557개소, 공공부문은 292개소이다. 1차 조사에서 특별조사로 실시한 공공부문 148개 기관을 제외하면 1차 조사부터 5차 조사까지 최종 조사된 모든 사업체의 수는 2,701개소이고, 이 중 민간부문은 2,557개소, 공공부문은 144개소이다.

### 제3절 6차년도 조사의 표본추가를 위한 표본설계 및 응답현황

#### 1. 표본추가를 위한 이중추출틀 방식의 표본설계

##### 가. 이중틀 표본설계 개요

사업체패널조사는 2006년 패널표본이 개편된 이후 격년 주기로 2014년까지 총 5차례에 걸쳐 진행되었다. 총 네 번의 추적조사를 시행하여 패널 이탈과 표본대체 등의 이유로 악화되어온 표본의 대표성과 지속가능성을 향상하기 위해, 신규 표본사업체의 추가 시 기존 패널모집단의 특성은 물론 표본추가 시점을 기준으로 횡단면모집단의 특성을 모두 반영할 수 있는 이중틀 방안이 검토되었다.

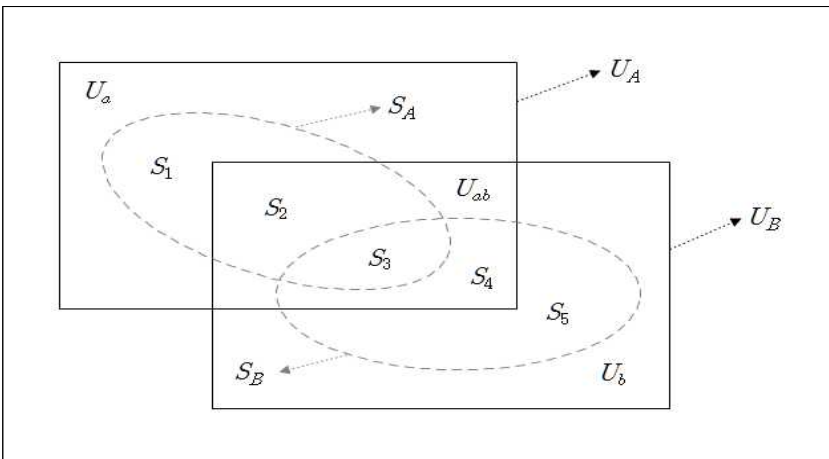
[그림 2-1]은 기존 표본의 (횡단면)모집단과 6차년도 횡단면모집단을  $U_A$ 와  $U_B$ 로 표기하고, 추적기간 동안 적격성을 유지한 개체와 소멸한 개체를  $U_{ab}$ 와  $U_a$ 으로, 5차년도 이후 생성된 개체를  $U_b$ 로 표기하고 있다. 기존 표본의 횡단면모집단은  $U_A = U_a \cup U_{ab}$ 이고 신규표본의 횡단면모집단은  $U_B = U_b \cup U_{ab}$ 으로 표기한다. 더불어 기존 표본의 횡단면모집단은 휴·폐업에 따른 소멸분  $U_a$ 와 적격성 유지분  $U_{ab}$ 로 나뉘며, 후자의

경우에는 구성 및 특성 변화를 파악할 수 있는 패널조사의 대상이 된다.

기존표본과 신규표본의 횡단면모집단  $U_A$ 와  $U_B$ 으로부터 추출된 표본은 각각  $s_A$ 와  $s_B$ 가 되며, 두 표본에 중복된 업체  $s_3$ , 휴·폐업에 의한 소멸분  $s_1$ , 신생업체  $s_5$ , 중복되지 않고 시점에 존재하는 표본  $s_2$ 와  $s_4$ 로 나누어 구분할 수 있다. 두 시점에 걸쳐 존재하는 패널모집단  $U_{ab}$ 에 속하는 표본은  $s_{ab}$ 로 이는  $s_2, s_3, s_4$ 로 구성된다. 이후 신규표본의 표본설계에 대한 논의는 [그림 2-1]의 모집단 분해도를 기준으로 한다.

이중틀 방식은 일반적으로 (고정된 시점의 특정) 모집단에 대해 서로 다른 포함률을 갖는 두 개의 표본추출틀을 함께 사용하는 것을 대상으로 고려된다. 하지만 본 연구에서는 생성, 소멸, 분리, 병합, 규모 및 분류 변경 등과 같은 모집단 변화가 반영된 사업체 명부가 관리되지 않고 시점별 명부만 존재하는 상황에서 제한적이거나 모집단 변동을 파악하고자 하는 노력의 일환으로 이중틀 접근을 고려하였다. 따라서 1~5차 조사를 위해 추출되었지만 무응답과 소멸이 아닌 분리, 병합, 규모 및 분류변경 등의 이유로 패널에서 이탈되었던 개체들은 추가표본에 포함될 수 있는 가능성이 존재함에 유의하여야 할 것이다.

[그림 2-1] 목표모집단 및 표본 분해도



### 나. 표본추가를 위한 표본추출틀

6차년도 사업체패널조사의 횡단면모집단에 대한 표본추출틀은 2014년 12월 말 기준 『전국사업체조사』의 사업체 명부를 토대로 한다. 여기에서 (금융·보험업, 전기·가스·수도업 등의) 일부 업종을 제외한 사업체들이 사업체패널조사 및 표본추출의 단위가 되며, 이 중 조사대상은 상시근로자 30인 이상인 사업체 가운데 조사수행, 자료분석 등을 고려하여 일부 분류를 추가로 제외한 총 57,894개<sup>8)</sup> 사업체이다. 단, 고용보험 DB와의 연계에 대해서는 고려하지 않았다(박인호·김기민, 2016).

〈표 2-7〉 업종분류

	대분류	중분류	산업분류코드
1	제조업	경공업	10~18, 32, 33
2		화학공업	19~23
3		금속, 자동차, 운송	24, 25, 29~31,
4		전기, 전자, 정밀	26~28
5	비제조업	건설업	41, 42
6		전기, 가스, 수도업	35, 36
7		개인서비스업	37~39, 45~47, 55, 56, 59, 60, 90~98
8		유통서비스업	49~52, 61
9		사업서비스업	58, 62, 63, 64~75,
10		사회서비스업	84, 85, 86, 87, 99

모집단 층화는 산업분류와 사업체 규모를 각각 10개와 4개로 구분하여 최대 40개 층으로 구성하였다. 산업분류는 제조업, 건설업, 전기·가스·수도업의 경우 과거 사업체패널조사의 분류와 동일하게 정하였고, 서비스업의 경우에는 OECD 분류기준을 적용하여 4가지로 분류<sup>9)</sup>하였다. <표 2-7>

8) 이 중 사업체패널조사에 포함되어 있는 기존 사업체 수는 총 1,451개소이다.

9) OECD에서 정의하고 있는 4가지 범주는 사업서비스업(Producer service), 유통(배분)서비스업(Distributive service), 개인서비스업(Personal service)과 사회서비스업(Social service)로 구성된다. 사업서비스업은 금융 및 보험업, 부동산 및 임대업, 사업서비스업을 포함하고 유통(배분)서비스업은 운수업, 통신업을 포함하며, 개인

은 모집단 층화에 고려된 10가지 업종분류에 대해 기술하고 있다. 사업체 규모는 상시근로자, 즉 상용근로자와 임시·일용 근로자 기준으로 30~99인, 100~299인, 300~499인, 500인 이상의 4가지 분류를 고려하였다.

<표 2-8>은 조사대상으로 포함되는 사업체, 즉 조사모집단의 산업분류와 사업체 규모별 분포를 사업체 수에 대해 각각 요약하고 있다.

<표 2-8> 조사모집단의 산업×규모별 사업체 수

(단위: 개소)

		30~99인	100~299인	300~499인	500인 이상	합 계
제조업	경공업	3,862	681	59	30	4,632
	화학공업	2,962	617	67	64	3,710
	금속·자동차·운송	7,401	1,494	166	118	9,179
	전기·전자·정밀	2,835	718	89	106	3,748
비제조업	건설업	3,729	803	115	92	4,739
	전기·가스·수도업	46	26	7	2	81
	개인서비스업	8,725	1,297	120	71	10,213
	유통서비스업	2,843	1,116	73	48	4,080
	사업서비스업	8,833	2,475	525	517	12,350
	사회서비스업	4,087	885	72	118	5,162
계		45,323	10,112	1,293	1,166	57,894

#### 다. 신규표본 추출결과

2016년도 추가표본을 위한 표본추출에서는 기존의 표본설계에서 고려하였던 지역배분은 제외하고 산업별 배분과 규모별 배분을 순차적으로 고려하되 기존 경험을 바탕으로 예측응답률을 고려하여 최종적인 표본배분을 수행하였다. 요약하여 정리하면 다음과 같다(박인호·김기민, 2016).

##### i) 업종별로 배분하되 전기·가스·수도업은 별도 추출(과거와 마찬가지로)

서비스업은 도·소매, 숙박 및 음식점업, 오락·문화 및 운동 관련업, 가사 서비스업, 기타 공공·수리 및 개인서비스업이며, 사회서비스업은 공공행정, 국방 및 사회보장행정, 보건 및 사회복지사업, 교육서비스업, 국제 및 외국 기관으로 구성된다(8차 한국표준산업분류에 의거).



가지로 25개 고정<sup>10)</sup>)

- ii) 업종 내 500인 이상 규모층에 대해 전수조사(약 39% 응답률<sup>11)</sup>) 가정
- iii) 500인 미만 규모층에 대해 표본조사 적용
- iv) 표본 제공 방식은 예측응답률<sup>12)</sup>을 고려하여 결정

표본배분은 산업분류와 사업체규모를 순차적으로 고려하였고 기본적으로 비례배분과 등배분의 절충형태인 키쉬배분(Kish allocation)을 적용하였다.

신규표본은 실질적으로는 두 차례에 걸쳐 추출되었다. 이는 조사환경 악화에 따라 기존 조사의 응답률을 적용한 앞서의 표본배분으로는 목표 응답수를 채울 수 없어서 2차 신규표본이 추가된 것이 2차 신규표본은 Keyfitz 방식(Brick et al., 1987)의 조건부 추출확률이 적용되었고 결과적

〈표 2-9〉 1차 신규표본 추가분

(단위: 개소)

		30~99인	100~299인	300~499인	합 계
제조업	경공업	154	294	-	448
	화학공업	244	155	-	339
	금속·자동차·운송	867	569	58	1,494
	전기·전자·정밀	438	449	-	887
비제조업	건설업	1,010	658	-	1,668
	개인서비스업	870	295	-	1,165
	유통서비스업	284	162	-	446
	사업서비스업	522	721	339	1,582
	사회서비스업	165	176	-	341
	전기·가스·수도업	18	13	-	31
계		4,72	3,492	397	8,461

- 10) 과거 25개 기준에 대한 이유를 찾지 못하여 해당 업종의 주요한 추정량들의 정도 수준(RSE 등)을 산출하여 판단하였다.
- 11) 2006년에 실시한 사업체패널조사에서 500인 이상 규모의 사업체 평균응답률이 약 39%로 나타남. 여기에 0.8을 곱하여 보수적 수치를 적용하였다.
- 12) 예측응답률은 2006년 조사경험을 토대로 추정하였고, 여기에 0.8을 곱하여 보수적 수치를 적용하였다.

으로 무조건부 추출확률이 층별로 동일할 수 있도록 하였다. <표 2-9>, <표 2-10>, <표 2-11>은 각각 1차 신규표본, 2차 신규표본 그리고 전체 신규표본 규모를 산업분류 및 사업체규모별로 나타내주고 있다.

<표 2-10> 2차 신규표본 추가분

(단위: 개소)

		30~99인	100~299인	300~499인	합 계
제조업	경공업	154	258	-	412
	화학공업	244	155	-	399
	금속·자동차·운송	867	589	-	1,436
	전기·전자·정밀	438	134	-	572
비제조업	건설업	1,010	-	-	1,010
	개인서비스업	870	294	-	1,165
	유통서비스업	284	162	-	446
	사업서비스업	522	721	-	1,243
	사회서비스업	165	176	-	341
	전기·가스·수도업	15	-	-	15
계		4,589	2,470	-	7,039

<표 2-11> 전체 신규표본 추가

(단위: 개소)

		30~99인	100~299인	300~499인	합 계
제조업	경공업	308	552	-	860
	화학공업	488	310	-	738
	금속·자동차·운송	1,734	1,158	58	2,930
	전기·전자·정밀	876	583	-	1,459
비제조업	건설업	2,020	658	-	2,678
	개인서비스업	1,740	589	-	2,330
	유통서비스업	568	324	-	892
	사업서비스업	1,044	1,442	339	2,825
	사회서비스업	330	352	-	682
	전기·가스·수도업	33	13	-	46
계		9,141	5,962	397	15,550

## 2. 응답현황

### 가. 응답현황 개요

<표 2-12>는 6차년도 조사에서 표본구분 및 응답유형별 표본사업체의 현황을 정리하고 있다. 기존 패널표본은 총 1,669개 사업장에 포함되었는데, 이 중 6차년도 조사기준으로 소멸된 사업장  $S_1$ 은 50개소, 신규표본과 중복되지 않는 사업장  $S_2$ 는 3건의 중복보고를 포함하여 총 899개소, 신규표본과 중복된 사업장  $S_3$ 은 총 720개소이다. 신규표본은 총 13,831개소로 이 중 패널표본과 중복되지 않은 사업장 가운데 2013년 이전에 설립된 사업장  $S_4$ 는 총 13,304개소<sup>13)</sup>이고 2014년에 설립된 사업장  $S_5$ 는 총 527개소이다.

<표 2-13>은 1차에서 6차 사업체패널조사의 전체 표본사업체 응답유형 현황을 보여준다. 6차 조사에서는 민간 응답사업체 2,136개가 추가되었다.

<표 2-12> 제6차년도 표본종류별 응답현황

(단위: 개소, %)

표본구분		적 합				부적합		합 계 (T)
		응답 (A)	무응답 (B)	소계 (C=A+B)	응답률 (A/C)	표본 사업체 (D)	부적합 비율 (D/T)	
기 존	$S_1$	-	-	-	-	50	100.0%	50
	$S_2$	709	139	848	83.6%	51	5.7%	899
	$S_3$	534	166	700	76.3%	20	2.8%	720
	소 계	1,243	305	1,548	80.3%	121	7.2%	1,669
추 가	$S_4$	2,076	9,911	11,987	17.3%	1,317	9.9%	13,304
	$S_5$	60	390	450	13.3%	77	14.6%	527
	소 계	2,136	10,301	12,437	17.2%	1,394	10.1%	13,831
합 계		3,379	10,606	13,985	24.2%	1,515	9.8%	15,500

13) 패널표본 중 조사에 참여하고 있지 않은 표본사업장일 수는 있으나 6차년도 조사에서 확인이 불가한 사업장을 포함할 수도 있다.

〈표 2-13〉 사업체패널조사 1~6차 전체 사업체에 대한 응답유형 현황

(단위: 개소, %)

	1차	2차	3차	4차	5차	6차	민간	공공	전 체
1	○	○	○	○	○	○	596(16.2)	85(71.4)	681(17.9)
2	○	○	○	○	○	×	203( 5.5)	14(11.8)	217( 5.7)
3	○	○	×	○	○	○	16( 0.4)	7( 5.9)	23( 0.6)
4	○	○	×	○	○	×	9( 0.2)	--	9( 0.2)
5	○	○	×	×	○	○	1( 0.0)	--	1( 0.0)
6	○	×	○	○	○	○	34( 0.9)	2( 1.7)	36( 0.9)
7	○	×	○	○	○	×	7( 0.2)	2( 1.7)	9( 0.2)
8	○	×	×	×	○	○	1( 0.0)	4( 3.4)	5( 0.1)
9	○	×	×	×	○	×	2( 0.1)	3( 2.5)	5( 0.1)
10		○	○	○	○	○	139( 3.8)	--	139( 3.7)
11		○	○	○	○	×	36( 1.0)	--	36( 0.9)
12			○	○	○	○	75( 2.0)	--	75( 2.0)
13			○	○	○	×	23( 0.6)	--	23( 0.6)
14				○	○	○	160( 4.3)	--	160( 4.2)
15				○	○	×	40( 1.1)	--	40( 1.1)
16					○	○	173( 4.7)	2( 1.7)	175( 4.6)
17					○	×	32( 0.9)	--	32( 0.8)
18						○	2,136(58.0)	--	2,136(56.2)
합계							3,683	119	3,802

## 나. 제6차 기준표본 응답현황

〈표 2-14〉는 표본종류별 횡단면 조사기준의 응답현황을 정리하고 있다. 횡단면조사의 적격 기준에서 상시근로자 25~29인은 30인 이상 사업체에 포함되지만 24인 이하는 가중치 산정과 분석 대상에서 제외하는 것이 기본 원칙이다(이기재, 2015: 23). 이에 따라 52개<sup>14)</sup> 사업체가 부적격한 것으로 제외되었다. 〈표 2-15〉는 제6차년도 사업체패널조사에서 횡단면기준치의

14) 상시근로자 25인 미만으로 (추가적) 부적격으로 분류된 사업체는 총 52개로  $S_2$ 와  $S_3$ 에서 각각 48개와 4개 사업체이다.

산출 대상 표본사업체와 응답사업체를, <표 2-16>은 1차년도 표본설계 당시의 산업대분류와 사업체 규모에 따른 현황을 각각 정리해 보여준다.

<표 2-14> 제6차년도 표본종류별 횡단면조사 기준 응답현황

(단위: 개소, %)

표본구분	적합				부적합		합계 (T)
	응답 (A)	무응답 (B)	소계 (C=A+B)	응답률 (A/C)	표본 사업체 (D)	부적합 비율 (D/T)	
$S_A$	1,243	305	1,548	80.3%	118	7.1%	1,666
$S_1$	-	-	-	-	50	100.0%	50
$S_2$	709	139	848	83.6%	48	5.4%	896
$S_3$	534	166	700	76.3%	20	2.8%	720

<표 2-15> 제6차년도 기존표본 횡단면조사 산출 대상 표본사업체 현황

(단위: 개소)

산업분류	30~99인	100~299인	300~499인	500인 이상	계
경공업	88	51	38	23	200
화학공업	79	31	28	23	161
금속, 자동차, 운송	96	40	44	32	212
전기, 전자, 정밀공업	73	34	27	22	156
건설업	66	31	6	3	106
개인서비스업	73	39	30	14	156
유통서비스업, 운수업	70	51	26	11	158
통신업	14	8	2	3	27
금융, 보험업	22	15	8	6	51
사업서비스업, 기타	60	41	40	33	174
사회서비스업	46	27	25	29	127
전기, 가스 및 수도사업	13	6	-	-	19
소 계	700	374	274	199	1,547
공공분야	-	-	-	-	119
계	700	374	274	199	1,666

〈표 2-16〉 제6차년도 기존표본 횡단면조사 기준 최종 응답현황

(단위: 개소)

산업분류	30~99인	100~299인	300~499인	500인 이상	계
경공업	63	46	31	17	157
화학공업	59	25	19	13	116
금속, 자동차, 운송	77	30	31	20	158
전기, 전자, 정밀공업	53	27	15	11	106
건설업	48	20	6	1	75
개인서비스업	57	34	22	6	119
유통서비스업, 운수업	54	43	23	9	129
통신업	6	4	-	1	11
금융, 보험업	17	13	6	5	41
사업서비스업, 기타	42	31	22	26	121
사회서비스업	34	22	19	17	92
전기, 가스 및 수도사업	13	6	-	-	19
소 계	523	301	194	126	1,144
공공분야	-	-	-	-	99
계	523	301	194	126	1,243

#### 다. 6차 조사 신규표본 응답현황

제6차 사업체패널조사를 위해 고려된 기존표본은 총 1,669개 사업체이며 이중  $S_2$ 의 3개 사업체가 중복 보고되어 이를 삭제하면 총 1,666개가 기존표본 횡단면가중치의 산출 대상이다. 신규표본은 기존표본 중  $S_3$  (720개)과 추가표본 중  $S_4$  (13,304개),  $S_5$  (527개)로 이루어지며 총 14,551개 사업체를 포함한다. <표 2-17>이 신규표본의 표본층인 산업분류 및 사업체 규모 분류에 따른 표본사업체 현황을 정리하고 있다. 공공분야는 총 48개이고 비공공분야는 총 14,503개 사업체로 이루어졌다. <표 2-18>은 신규표본의 산업분류별·사업체규모 분류별 응답사업체 현황을 정리하고 있다.

〈표 2-17〉 제6차년도 신규표본 횡단면조사 산출 대상 표본사업체 현황

(단위: 개소)

산업분류	30~99인	100~299인	300~499인	500인 이상	계
경공업	492	367	59	29	947
화학공업	494	251	67	63	875
금속, 자동차, 운송	1,271	676	166	118	2,231
전기,전자, 정밀공업	688	500	89	106	1,383
건설업	1,248	803	115	92	2,258
개인서비스업	1,377	428	120	70	1,995
유통서비스업	492	256	72	47	867
사업서비스업	1,109	864	510	507	2,990
사회서비스업	446	265	71	117	899
전기, 가스 및 수도사업	25	26	5	2	58
소 계	7,642	4,436	1,274	1,151	14,503
공공분야	6	10	19	13	48
계	7,648	4,446	1,293	1,164	14,551

〈표 2-18〉 제6차년도 신규표본 횡단면조사 산출 대상 응답사업체 현황

(단위: 개소)

산업분류	30~99인	100~299인	300~499인	500인 이상	계
경공업	141	88	31	7	267
화학공업	114	48	33	24	219
금속, 자동차, 운송	200	78	68	35	381
전기,전자, 정밀공업	110	60	41	14	225
건설업	116	71	28	17	232
개인서비스업	189	88	41	16	334
유통서비스업	98	73	41	6	218
사업서비스업	172	149	96	79	496
사회서비스업	120	48	35	30	233
전기, 가스 및 수도사업	11	12	3	-	26
소 계	1,271	715	417	228	2,631
공공분야	3	10	15	11	39
계	1,274	725	432	239	2,670

## 제4절 사업체패널조사 가중치 산정

### 1. 제6차 조사 종단면가중치

사업체패널조사의 조사모집단은 원칙적으로 1차년도 사업체패널조사의 조사기준 시점인 2005년도 12월 말 기준으로 상용근로자 30인 이상 전체 사업체이다. 따라서 추적조사에서 해당 사업체의 상용근로자수가 30인 미만이 되어도 조사 및 분석 대상이 된다.

패널가중치는 1차년도 기준 상용근로자 30인 이상 사업체를 대상으로 사업체의 전반적인 경영환경 및 인적자원관리체계, 노사관계의 현황 및 임금교섭 과정 등에 대한 정보를 이후 추적 조사하여 기간 동안 어떻게 변화하였는지를 파악하는 데 목적을 둔다.

패널사업체에 대해 기존 패널가중치 작성은 1차년도 조사의 횡단면 가중치를 기초가중치로 고려하였다. 이후 각각의 추적조사에서 응답성향에 대한 모형적합을 통해 패널이탈은 물론 무응답을 보정하여 패널가중치를 산출하였다. 예를 들면, 1차년도 조사에 응답한 표본사업장이 2차년도 추적조사에 응답할 확률은 1차년도 조사결과에 의해 파악한 사업장 특성변수를 설명변수로 하는 로지스틱회귀모형(logistic regression model)을 통해 추정하였다. 다시 말해 2차년도 패널조사에 대한 응답 여부를 반응변수로 하고, 산업분류, 지역, 근로자수, 노조 유무, 설립연도, 본·지사 등을 잠재적인 설명변수로 하는 로지스틱회귀모형을 이용하여, 개별 표본사업체의 응답확률(혹은 응답성향점수)을 추정하였다. 3차년도 이후의 추적조사는 1차 조사를 포함하여 관심 시점 바로 전 조사까지 모두 응답한 사업체들을 대상으로 하였다. 이렇게 추정된 표본사업체의 응답확률은 관심시점 추적조사에 응답한 사업체의 이전 시점 패널가중치로 나누어, 관심시점의 패널가중치를 다음과 같이 산출하였다.



$$pw_{tk} = \begin{cases} \frac{pw_{t-1,k}}{\hat{\rho}_{tk}} & \text{응답} \\ 0 & \text{기타} \end{cases} \quad (2-1)$$

여기서  $pw_{tk}$ 는 표본사업체  $k$ 의 추적시점  $t$ 의 패널가중치를 나타내며,  $\hat{\rho}_{tk}$ 는 로지스틱회귀모형에 적합하여 산출한 표본사업체  $k$ 가  $t$ -시점의 추적조사에 응답할 확률을 나타낸다. 패널구축을 위한 1차조사의 표본사업체 가중치는 해당 조사의 횡단면가중치  $pw_{1k} = w_k$ 가 된다.

6차 추적조사의 패널가중치 산출도 기존 패널가중치 산출방법과 동일하며 가중치 산출 대상은 1차에서 5차 조사까지 모두 응답한 총 898개<sup>15)</sup> 사업체들이다(표 2-6, 표 2-12 참조). <표 2-19>는 이들의 6차 추적조사의 응답 여부를 로지스틱회귀모형에 적합한 결과로 정리하고 있다. 패널조사의 응답률에 영향을 미치는 설명변수는 산업구분, 사업체 규모, 사업체 구분 순으로 유의하며, 지역구분은 유의한 설명력을 갖고 있지 않지만 5차 패널조사에서 응답성향모형에 산업구분과 사업체규모와 함께 포함되어 6차 패널조사의 응답성향모형에도 포함시켰다.

6차 조사의 패널가중치는 식 (2-1)과 같이 5차 패널가중치  $pw_{5k}$ 를 응답성향점수  $\hat{\rho}_{6k}$ 로 나누어 산출하였다. <표 2-20>은 5~6차 패널가중치의 분포를 비교하고 있다. 5차 조사에 비해 평균, 표준편차, 범위 등이 각각 증가하였지만 평균값의 변동에 비해 표준편차의 변동은 작아 변동계수는 오히려 약간 감소하였다.

15) 6차년도 조사부터 교육서비스 업종이 제외되었고 이에 따라 1~5차 패널표본에서도 제외되었다.

〈표 2-19〉 1~5차 사업체패널조사의 전체 사업체 응답유형 현황

(단위: 개소)

예측변수	수준	자유도	카이제곱	유의확률	추정값
절편					1.9130
산업구분	(경공업)	9	2,114.9729	<.0001	
	화학공업				0.1953
	금속/자동차/운송				0.9276
	전기/전자/정밀공업				-0.3241
	건설업				-0.3597
	개인서비스업				-0.0846
	유통서비스업/운수업/통신업				0.2364
	금융/보험/사업서비스				-0.2191
	사회서비스업				0.7651
	전기/가스/수도				12.4727
지역구분	(서울)	4	6.4537	0.1677	
	인천/경기				0.4288
	강원/대전/충북/충남				1.4969
	광주/전북/전남/제주				-0.1662
	부산/대구/울산/경북/경남				0.2776
근로자 수	(500인 이상)	5	21.2293	0.0007	
	30인 미만				0.3763
	30~49인				0.2021
	50~99인				-1.0957
	100~299인				-0.7134
	300~499인				-1.4380
사업체구분	(기타)	3	16.3728	0.0010	
	단독				-0.6158
	다수사업체-본사				0.9378
	다수사업체-본사 아님				-1.0065

〈표 2-20〉 사업체패널조사의 5~6차 패널가중치 분포

(단위: 개소)

	규모	평균	표준편차	변동계수	최소	25%	50%	75%	최대
5차	898	40.1	45.1	112.5	2.0	4.7	24.3	68.2	252.7
6차	681	49.5	51.7	104.4	2.5	5.8	32.0	83.3	311.7

## 2. 제6차 조사 횡단면가중치

### 가. 제6차 횡단면가중치 개요 및 표본구성

이중틀 접근에 의한 표본추가로 인해 기존표본  $s_A$ 와 신규표본  $s_B$ 로 구성되었고 이에 개별 표본별 횡단면가중치가 구축되었다. 기존표본과 신규표본의 횡단면 가중치 산정은 기존방식(이기재, 2015)과 동일하게 적용하였다. 최초년 혹은 2015년 표본 추출 시 사용하였던 산업대분류와 사업체 규모를 층으로 하는 복합표본설계(complex sample design)를 반영한 ㉠ 설계가중치, ㉡ 무응답 조정, ㉢ 사후층화 조정 등의 세 가지 구성요소를 통합하여 산정한다.

기존표본과 신규표본의 횡단면가중치는 의사최대가능도(擬似最大可能度, pseudo-maximum likelihood, 이후 PML) 추정(Skinner and Rao, 1996)에 근거한 가중치를 산출한다. 따라서 6차 조사의 횡단면가중치는 다음과 같은 세 종류로 이루어진다.

- ① 기존표본  $s_A (= s_1 \cup s_2 \cup s_3)$ 에 대한 횡단면가중치  $w_{Ak}$
- ② 신규표본  $s_B (= s_3 \cup s_4 \cup s_5)$ 에 대한 횡단면가중치  $w_{Bk}$
- ③ 전체표본  $s (= s_A \cup s_B)$ 에 대한 PML 횡단면가중치  $w_{PML,k}$

### 나. 기존 사업체패널표본 횡단면가중치

제6차 사업체패널조사를 위해 고려된 기존표본은 총 1,669개 사업체(그림 2-1, 표 2-12 참고)이며 이 중  $S_2$ 의 3개 사업체가 중복 보고되어 이를 제거하면 총 1,666개가 기존표본 횡단면가중치의 산출 대상이다.

기존표본  $s_A$ 는 기존모집단  $U_A$ 로부터 추출된 사업체이지만 3.1절에서 기술된 것처럼 1차년도(2006년)는 물론 2차년도(2008년), 3차년도(2010년), 4차년도(2012년), 5차년도(2014년)에 신생된 사업체들을 포함하고 있다(표 2-12 참고). 이중틀 접근의 일반적 형태와는 달리 기존조사에 대한 모집단은 2006년 이후 2014년까지의 생멸을 포함하는 사업체들로 구성된다. 이

는 패널이탈로 인한 지속적인 표본사업체 감소를 방지하기 위해 사업체패널조사의 횡단면적 추정이 가능하도록 만들고자 2차 조사부터 표본사업체 추가로 사업체패널을 보완해 왔기 때문이다(이기재, 2015).

사업체패널조사에서 횡단면가중치 산정의 대상은 원칙적으로 해당 조사 기준시점에서 상용근로자 30인 이상인 응답사업체이다. 또한 이전 차수에서 무응답한 사업체일지라도 관심 차수에 다시 참여한다면 응답업체로 분류한다. 하지만 조사 기준시점과 실제 조사시점 간 시차 발생은 물론 상용근로자 30인 이상 사업체로 국한할 경우에 응답사업체 정보 중 상당 부분을 활용할 수 없다는 점을 고려하여 상시근로자 25인 이상 표본사업체를 횡단면가중치 산출 대상으로 삼았다. 2~5차년도 사업체패널조사에서도 동일한 기준에 따라 횡단면 가중치 산출 대상을 국한하였다.

기존표본의 횡단면가중치는 기존방식(이기재, 2015)에 따라 설계가중치, 무응답 조정, 사후층화 조정이 순차적으로 적용되었다. 먼저, 2차년도 이후 추가된 표본사업체를 1차년도 표본설계에서의 해당 산업 및 규모 층에서 층화 임의 추출된 것으로 가정하여 설계가중치를 구하였다(이기재, 2015, 표 3).

다음 단계로 6차 조사에서 발생한 무응답에 대한 조정을 고려하였다. 여기서는 기존방식과 마찬가지로 일반 사업장들에 대해 1차 조사 시 표본설계의 산업분류와 사업장 규모를 이용하였고 공공부문에 대해서는 전체를 하나의 층으로 간주하여 무응답 조정을 수행하였다. 먼저, 민간부문에 대한 단위 무응답을 보정하기 위한 무응답 조정계수는 다음 식에 의해 산출하였다.

$$NR_{Ac} = \frac{\sum_{i \in U_c \cap s_{AXR}} cw_{Ai} + \sum_{i \in U_c \cap s_{AXNR}} cw_{Ai}}{\sum_{i \in U_c \cap s_{AXR}} cw_{Ai}}$$

여기서  $cw_{Ai}$ 는 기존표본  $s_A$ 의 설계가중치이고,  $U_c$ 는 무응답 조정층으로 산업대분류와 사업체 규모로 결정되며,  $s_{AXR}$ 과  $s_{AXNR}$ 는 각각 횡단면조사 기준의 응답사업체와 무응답사업체를 나타낸다.

무응답조정계수를 구한 후에 무응답조정가중치는 다음과 같이 설계가중치에 무응답조정계수의 곱으로 정의된다.

$$cw_{Ai}^{NR} = \begin{cases} cw_{Ai} \times NR_{Ac} & i \in s_{AXR} \\ 0 & i \in s_{AXNR} \end{cases}$$

반면, 공공부문의 경우에는 1차 조사 표본설계 당시의 모집단 사업체수 365개와 제6차 조사 시 응답사업체의 비인  $cw_{Ai}^{NR} = 365/100$ 의 값이 응답사업체에 대해서만 부여되었다.

사후층은 6차 조사에서 파악한 사업체의 실질적인 산업대분류와 사업체규모로 정의하였다. 공공부문은 기존표본에서만 포함되므로 무응답 조정에서와 마찬가지로 사후층화에서도 이를 반영하여 일반사업체에 대해서만 고려하였다.

#### 다. 신규 사업체패널표본 횡단면가중치

6차 조사를 위한 신규표본은 총 14,551개 사업체(그림 2-1, 표 2-12 참고)로 구성된다. 이중틀 추출방식에 따라 기존표본과 중복되어 추출된 표본  $S_3$ 는 720개의 사업체로 이루어지며,  $S_4$ 는 기존표본조사 시점에 존재하나 기존표본에 포함되지 않고 새롭게 추출된 13,304개 사업체로 이루어지고,  $S_5$ 는 2016년 조사시점을 기준으로 한 신생업체 527개 사업체로 구성된다. 따라서 6차 조사에서 순수하게 추가된 신규표본은  $S_4$ 와  $S_5$ 인 13,831개 사업체이다.

6차 조사의 횡단면가중치 산정의 대상은 기존표본과 동일하게 원칙적으로 조사 기준시점에서 상시근로자 25인 이상인 표본사업체를 횡단면가중치 산출의 대상으로 삼았다.

신규표본의 횡단면가중치는 기존표본과 동일하게 설계가중치, 무응답 조정, 사후층화 조정이 순차적으로 적용되었다. 먼저, 6차년도 표본설계에서의 해당 산업 및 규모 층에서 층화 임의 추출을 반영하여 다음과 같이 설계가중치를 구하였다.

$$cw_{Bi}^D = N_h/n_h$$

다음 단계로 6차 조사에서 발생한 무응답에 대한 조정을 고려하였는데 기존표본과 마찬가지로 일반 사업장들에 대해서는 산업대분류와 사업장

규모로 구성된 표본층 내에서 단위무응답을 보정하기 위한 무응답 조정 계수를 다음 식으로 산출하였다.

$$NR_{Bh} = \frac{\sum_{i \in U_{Bh} \cap s_{BXR}} cw_{Bi}^D + \sum_{i \in U_{Bh} \cap s_{BXNR}} cw_{Bi}^D}{\sum_{i \in U_{Bh} \cap s_{BXR}} cw_{Bi}^D}$$

여기서  $cw_{Bi}^D$ 는 신규표본  $S_B$ 의 설계가중치이고,  $U_{Bh}$ 는 표본층인 무응답조정층이며,  $s_{BXR}$ 과  $s_{BXNR}$ 는 각각 횡단면조사 기준의 응답사업체와 무응답사업체를 나타낸다.

무응답조정계수를 구한 후에 무응답조정 가중치는 다음과 같이 설계가중치에 무응답 조정계수의 곱으로 정의된다.

$$cw_{Bi}^{NR} = \begin{cases} cw_{Bi} \times NR_{Bh} & i \in s_{BXR} \\ 0 & i \in s_{BXNR} \end{cases}$$

사후층은 6차 조사에서 파악한 사업체의 실질적인 산업대분류와 사업체규모로 정의하였다. 산업대분류는 표본설계와 마찬가지로 10대 산업분류를 고려하였다. 사업체규모는 상시근로자수를 기준으로 화학공업, 금속자동차운송, 전기전자정밀, 사업서비스업, 사회서비스업에서는 총 6개 분류인 25~49인, 50~99인, 100~299인, 300~499인, 500~999인, 1,000인 이상을 적용하였다. 반면에 경공업, 건설업, 개인서비스업, 유통서비스업, 전기가스수도업의 5개 산업부문에 대해서는 25~49인, 50~99인, 100~299인, 300~499인, 500인 이상의 총 5개 분류를 적용하였다.

$$cw_{Bi}^{PS} = cw_{Bi}^{NR} PS_e$$

여기서 사후층화조정계수는 사후층  $e$ 의 모집단 사업체수  $N_e$ 와 표본가중합  $\hat{N}_e = \sum_{s_{BXR} \cap U_e} cw_{Bi}^{NR}$ 의 비로 다음과 같이 정의된다. <표 2-21>은 산업분류와 사업체규모로 구분한 사후층별 총합을 정리하고 있다.

$$PS_e = N_e / \hat{N}_e$$

〈표 2-21〉 사업체패널조사의 사후층 크기

(단위: 개소)

산업분류	사업체규모(상시근로자수)					계
	25~49인	50~99인	100~299인	300~499인	500인 이상 <sup>1)</sup>	총계
경공업	2,412	1,375	659	60	27	4,533
화학공업	1,849	1,137	617	70	(38, 21)	3,732
금속자동차운송	4,442	2,905	1,468	163	(69, 49)	9,096
전기전자정밀	1,625	1,109	690	89	(53, 43)	3,609
건설업	3,018	1,864	1,048	163	145	6,238
개인서비스업	6,018	3,421	1,453	140	81	11,113
유통서비스업	1,729	1,537	1,253	90	62	4,671
사업서비스업	5,568	4,071	2,721	597	(368, 160)	13,485
사회서비스업	2,302	2,075	939	72	(56, 62)	5,506
전기가스수도업	27	23	27	6	3	86
계	28,990	19,517	10,875	1,450	1,237	62,069

주: 1) 500인 이상의 열에서 ( , )는 (500~999인, 1,000인 이상)을 의미함.

#### 라. 이종틀 PML 횡단면가중치 산출

사업체 DB가 시점별로 연계되어 생성과 소멸, 업종 및 규모변화 등의 정보가 누적되며 사업체 고유번호로 갱신되어 관리된다면 사업체 생멸 및 변동이 더 잘 반영될 수 있다. 예를 들어, 미국 노동통계국(BLS, U.S. Bureau of Labor Statistics)은 표본틀의 재정비는 물론 주요 조사들의 재설계를 통해 조사들 간의 표본 할당 및 추출을 중앙집중식으로 처리한다. 이때 신규업체에 대해서는 독립적으로 영구난수(permanent random number)를 발생시키고, 표본층별로 조정하여 기존 표본에 추가하는 한편 기존 표본 중에서 폐업한 사업체는 제거하는 등 사업체 모집단의 변동을 반영하고 있다(Butani et al., 1998).

하지만 우리나라에서는 현재 사업체 DB가 마련되어 있지 않아 시점간 사업체 변동에 대한 내역을 파악할 수 없다. 이러한 상황에서 이중추출틀 방식을 응용한 표본추출은 기존표본에 더불어 최신 사업체 명부를 이용

한 신규표본을 추출함으로써 간접적이고 제한적이거나 시점별 모집단 변동을 파악할 수 있게 한다. 이중추출틀 방식은 두 개의 서로 다른 추출틀로부터 독립적으로 표본을 추출하므로 공통부분으로 추출된 표본에 대해서는 양쪽에서 추출될 수 있는 이중적인 추출확률이 감안되어 가중치가 절충적으로 산출되어야 한다. 본 연구에서는 Skinner and Rao(1996)의 의사최대가능도 가중치 방식을 적용하였다.

기존표본은 개념적으로 2006년부터 2014년까지의 횡단면 모집단으로부터 추출되었다고 가정할 수 있다. 하지만 실질적으로는 6차 조사 시점에서 현재 2006년부터 2014년까지 적격성을 유지하는 사업체  $U_{ab}$ 는 물론 이미 폐업한 사업체  $U_a$ 도 포함한다. 따라서 응답업체만으로는 기존패널 중 폐업한 사업체표본  $s_1$ 을 제외한 표본  $S_{ab}' = S_2 \cup S_3$ 만을 파악할 수 있고 해당 표본수는  $n_{ab}' = n_2 + n_3$ 이다. 표본  $A$ 의 최종가중치  $cw_{Ai}^{PS}$ 를 이용한 모집단 총수 추정량은  $\hat{N}_{ab}' = \sum_{i \in S_{ab}'} cw_{Ai}^{PS}$ 로 계산할 수 있다.

반면 신규표본은 2006년부터 2014년까지 적격성을 유지하는 사업체  $U_{ab}$ 는 물론 2014년 이후 생성된 신규사업체  $U_b$ 로부터 추출된다. 따라서 신규표본은  $S_B = S_{ab}'' \cup S_b$ 로 2006년부터 2014년까지 적격성을 유지한 크기  $n_{ab}'' = n_3 + n_4$ 의 표본  $S_{ab}'' = S_3 \cup S_4$ 와 2014년 이후 생성된 크기  $n_b = n_5$ 의 신생업체 표본  $S_b = S_5$ 로 분리된다. 신규표본의 전체 크기는  $n_B = n_{ab}'' + n_b$ 이다. 표본  $B$ 의 최종가중치  $cw_{Bi}^{PS}$ 를 이용한 모집단  $U_B$ ,  $U_{ab}$ ,  $U_b$ 의 크기는 각각  $\hat{N}_B = \hat{N}_{ab}'' + \hat{N}_b$ 과  $\hat{N}_{ab}'' = \sum_{i \in S_{ab}''} cw_{Bi}^{PS}$ , 그리고  $\hat{N}_b = \sum_{i \in S_b} cw_{Bi}^{PS}$ 로 계산할 수 있다.

따라서 동일한 모집단  $U_{ab}$ 에 대해 두 가지 표본  $S_{ab}'$ 과 표본  $S_{ab}''$ 로부터 추정되는 모집단 크기 추정량과 표본수를 이용하여 다음과 같이 절충된 PML 횡단면가중치를 산출할 수 있다.

$$\hat{cw}_i' = \left( \frac{\tilde{N}_{ab, PML}}{\hat{N}_{ab}'} \right) \left( \frac{n_{ab}'}{n_{ab}' + n_{ab}''} \right) cw_{Ai}^{PS}, \quad i \in S_{ab}'$$



$$\widehat{cw}_i'' = \begin{cases} \left( \frac{\widehat{N}_B - \widetilde{N}_{ab,PML}}{\widehat{N}_b} \right) cw_{Bi}^{PS} & i \in S_b \\ \left( \frac{\widetilde{N}_{ab,PML}}{\widehat{N}_{ab}''} \right) \left( \frac{n_{ab}''}{n_{ab}' + n_{ab}''} \right) cw_{Bi}^{PS} & i \in S_{ab}'' \end{cases}$$

여기서  $\widetilde{N}_{ab,PML} = (2p)^{-1} [q - (q^2 - 4pr)^{0.5}]$ 은  $U_{ab}$  규모에 대한 의사 최대가능도 추정량이며,  $r = n_{ab}' \widehat{N}_{ab}' \widehat{N}_B + n_B \widehat{N}_{ab}'' \widehat{N}_{ab}''$ ,  $p = n_{ab} + n_B$ ,  $q = n_{ab}' \widehat{N}_B + n_B \widehat{N}_{ab}' + n_{ab}' \widehat{N}_{ab}' + n_B \widehat{N}_{ab}''$ 이고  $\widetilde{n}_{ab}'' = n_B \widehat{N}_{ab}'' / \widehat{N}_B$ 이다.

의사최대가능도 가중치는 기존표본과 신규표본의 표본가중치를 절충한 다음의 형태로 정의된다.

$$cw_i^{PML} = \begin{cases} \widehat{cw}_i' & i \in S_1 \cup S_2 \\ \lambda \widehat{cw}_i' + (1 - \lambda) \widehat{cw}_i'' & i \in S_3 \\ \widehat{cw}_i'' & i \in S_4 \cup S_5 \end{cases}$$

여기서  $\lambda = cv_w''^2 / (cv_w'^2 + cv_w''^2)$ 은 표본 간 절충계수에 해당하며,  $cv_w'$ 는 산업별 기존표본 가중치  $\widehat{cw}_i'$ 와의 상대표준오차와  $cv_w''$ 는 산업별 신규표본 가중치  $\widehat{cw}_i''$ 의 상대표준오차를 각각 나타낸다.

〈표 2-22〉 이중틀 표본종류와 산업분류별 응답사업체 현황

(단위: 개소)

산업분류	표본종류					합 계
	$S_1$	$S_2$	$S_3$	$S_4$	$S_5$	
경공업	-	84	55	209	3	351
화학공업	-	66	41	176	2	285
금속자동차운송	-	93	79	293	9	474
전기전자정밀	-	44	47	173	5	269
건설업	-	29	42	186	4	261
개인서비스업	-	103	45	273	16	437
유통서비스업	-	88	47	170	1	306
사업서비스업	-	100	91	398	7	596
사회서비스업	-	35	37	184	12	268
전기가스수도업	-	7	11	14	1	33
합 계	-	649	495	2,076	60	3,280

〈표 2-23〉 PML 횡단면가중치 산출 구성요소

(단위 : 개소)

산업분류	$\hat{N}_{ab}'$	$\hat{N}_{ab}''$	$\hat{N}_b$	$\hat{N}_B$	$n_{ab}'$	$n_{ab}''$	$n_B$	$\tilde{N}_{ab,PML}$	$\gamma_{ab}'$	$\gamma_{ab}''$	$\gamma_b$	$\lambda$
경공업	3,916.0	4,415.0	85.9	4,497.8	139	264	267	3,911.6	0.3742	0.6258	7.0188	0.3739
화학공업	3,897.2	3,580.5	30.0	3,600.6	107	217	219	3,580.5	0.3110	0.6890	0.6718	0.4212
금속자동차운송	9,042.4	8,898.6	211.6	9,044.3	172	372	381	8,898.6	0.3111	0.6889	0.6890	0.4089
전기전자정밀	3,288.6	3,358.1	74.7	3,432.8	91	220	225	3,288.6	0.2969	0.7031	1.9306	0.4497
건설업	6,514.7	6,291.5	119.0	6,410.6	71	228	232	6,319.4	0.2314	0.7686	0.7657	0.4401
개인서비스업	10,288.3	10,448.5	730.4	11,178.9	148	318	334	10,288.3	0.3250	0.6750	1.2193	0.4359
유통서비스업	4,632.36	4,596.0	34.7	4,630.6	135	217	218	4,609.2	0.3824	0.6177	0.6176	0.5263
사업서비스업	12,882.0	13,422.6	276.7	13,699.3	191	489	496	12,820.0	0.2905	0.7095	2.9541	0.3898
사회서비스업	5,661.6	5,164.3	341.7	5,506.0	72	221	233	5,245.0	0.2311	0.7689	0.7639	0.1286
전기가스수도업	58.4	79.1	3.9	83.0	18	25	26	58.4	0.4961	0.5039	6.3874	0.5000

이중틀(dual frame) 의사최대가능도 가중치의 산출은 안정성 확보를 위해 제6차 신규표본 추출을 위한 표본설계 시 고려된 산업대분류를 기준으로 하며 당시 표본설계 기준의 산업분류가 없는 106개 사업체는 6차 조사기준 산업분류로 대체하였다.

<표 2-22>는 이중틀 분해에 따른 표본종류와 산업분류별 응답사업체의 현황을 나타낸다. 휴폐업에 따른 소멸 부분인  $S_1$ 은 정의상 응답사업체는 존재하지 않으며 응답 신생업체  $S_5$ 는 60개소이며, 생멸 부분이 아닌 기존표본과 신규표본에 중복되는 응답표본  $S_3$ 은 495개 업체이고 기존표본과 신규표본에만 속하는 응답표본인  $S_2$ 와  $S_4$ 는 각각 649개와 2,076개소로 구성되었다. <표 2-23>은 산업분류별 PML 횡단면가중치 산출에 필요한 구성요소들을 정리하여 주고 있다.

#### 마. 최종 횡단면가중치 산출

최종적으로 산업분류, 사업체 규모, 노동조합 여부에 따른 표본가중합이 모집단 사업체수와 같아지도록 PML 횡단면가중치를 조정하였다. 하지만 산업분류·사업체 규모와 함께 노동조합 여부를 고려하게 됨으로써 모집단이 지나치게 상세히 분류되어 사후층화 방법은 매우 불안정할 수 있다. 따라서 대안적 접근으로 산업분류와 노동조합 여부, 그리고 사업체 규모와 노동조합 여부의 조합 구분에 기초한 각각의 주변합을 이용해 레이킹-비 조정(raking-ratio adjustment)을 고려함으로써 최종 횡단면가중치를 산출하였다. 개별 사업체의 레이킹-비 조정 특성분류를 나타내는 지시자  $x_i$ 이라면 레이킹-비 조정 가중치는 다음과 같이 산출된다.

$$cw_i^F = cw_i^{PML} \exp(\tau' x_i)$$

이때  $cw_i^{PML}$ 은 PML 횡단면가중치,  $cw_i^F$ 는 레이킹-비 조정이 된 최종 횡단면가중치,  $\tau$ 는 다음의 칼리브레이션 조건을 만족시키는 해벡터(solution vector)가 된다.

$$\sum_{i \in s_{XR}} cw_i^F x_i = \sum_{i \in U_X} x_i$$

여기서  $s_{XR}$ 과  $U_X$ 는 횡단면 응답표본과 모집단을 각각 나타낸다.

<표 2-24>와 <표 2-25>는 앞서의 레이킹-비 조정에 사용된 모집단 총합정보를 포함하며, 산업분류 및 노동조합 유무 여부 그리고 사업체규모 및 노동조합 유무 여부에 따른 모집단 사업체 분포를 각각 보여준다. 총합정보는 고용노동부의 임금구조조사와 통계청의 2015년 전국사업체조사 자료를 이용하여 구성하였다.

<표 2-24> 산업분류 및 노동조합 여부에 따른 사업체 분포

(단위: 개소)

산업분류	노동조합		합계
	유	무	
경공업	640	3,893	4,533
화학공업	546	3,186	3,732
금속자동차운송	764	8,332	9,096
전기전자정밀	818	2,791	3,609
건설업	407	5,832	6,239
개인서비스업	1,877	9,235	11,112
유통서비스업	3,484	1,187	4,671
사업서비스업	2,210	11,293	13,503
사회서비스업	730	4,776	5,506
전기가스수도업	58	28	86
전 체	11,534	50,553	62,087

<표 2-25> 사업체규모 및 노동조합 여부에 따른 분포

(단위: 개소)

사업체 규모	노동조합		전 체
	유	무	
30~99인	6,917	41,590	48,507
100~299인	3,406	7,467	10,875
300~499인	541	909	1,450
500인 이상	670	584	1,254
전 체	11,534	50,553	62,087

<표 2-26>은 5차 조사 횡단면가중치를 포함하여 기존표본 및 신규표본, 그리고 통합표본의 횡단면가중치를 비교하고 있다. 6차 조사를 위한 표본추가로 인해 5차 횡단면가중치와 6차 기존표본 및 신규표본의 가중치에 비해 통합표본의 PML 횡단면가중치와 레이킹-비 가중치의 변동이 아주 작은 것을 확인할 수 있다.

<표 2-26> 사업체패널조사의 제6차 횡단면 가중치 분포

(단위: 개소)

	응답 수	평균	표준 편차	변동 계수	최소	25%	50%	75%	최대
5차 표본	1,484	30.6	34.2	111.9	0.4	6.1	20.0	39.2	233.5
6차 기존표본	1,144	52.6	65.4	124.4	0.6	8.1	31.6	70.4	499.9
6차 신규표본	2,136	41.0	41.0	144.2	1.0	3.9	11.4	35.5	388.2
6차(PML) 통합표본	3,280	18.0	17.0	94.8	0.1	4.8	15.3	27.7	194.8
6차(레이킹-비 조정)	3,280	18.9	17.6	93.5	0.2	6.4	15.0	26.0	268.3
전체(공공포함)	3,380	18.5	17.6	95.4	0.2	5.3	14.6	26.0	268.3

## 제3장

## 웹을 활용한 조사 효과에 대한 실증연구 II :

노동패널 CAWI 1차, 2차 실험 연구 결과 분석<sup>16)</sup>

## 제1절 서론

노동패널 자료는 동일한 대상에 대해 일정한 기간 동안 추적 반복 측정하여 그 변화를 관찰하는 것이다. 신뢰성 있는 패널자료를 생산하기 위해 일차적으로 대표성 있는 최초 표본을 선정하였고, 매해 조사에서는 가급적 많은 표본을 유지하도록 노력해왔다. 그럼에도 불구하고, 개인의 사망이나 이민, 이사 후 주소 추적 불가, 해를 거듭한 조사에 피로를 느낀 응답자의 응답 거부 등의 이유로 표본이탈(attrition)은 필연적으로 발생할 수밖에 없었다. 최근에는 이사 추적 실패의 사유가 점점 많아지면서, 현재의 대면조사 기반의 조사 방법과 패널 관리 방법에 한계를 느끼고 있다. 이에 노동패널팀에서는 현재 진행되고 있는 CAPI(Computer Assisted Personal Interviewing) 이외의 대안으로서 웹을 기반으로 한 CAWI(Computer Assisted Web Interviewing) 도입을 검토 중에 있다. 본 조사에 본격적으로 도입하기에 앞서 2016년도 노동패널 응답자와 별도로 표

16) 본 장은 2017년 이루어진 1차 웹조사 실험 연구 보고서인 김유빈 외(2017), 『패널자료 품질개선 연구(VII)』 제7장 『웹을 활용한 조사의 효과에 대한 실증연구: 노동패널 CAWI 실험 데이터의 분석』(이지은·신선옥·이혜정·김기홍 저술)의 후속 연구로 이루어진 것이다.

본을 설계하여 CAWI 실험 연구를 실행하였다.

1차 CAWI 실험의 목적은 ① 면접원이 없어도 응답자가 설문을 잘 이해하고 응답하는가? ② 조사방법에 따라 설문응답의 일관성이 유지되는가? 즉 주요 문항들에서 CAPI와 CAWI의 응답값이 차이가 나지 않는가? ③ 조사도구의 변화에도 현재의 조사성공률을 유지할 수 있을까를 살펴보는 것이었다.

이상의 목적에 따라 1차 보고서에서는 조사도구별, 즉 CAWI와 CAPI 조사 방법에 따른 항목별 결과의 차이를 검증하였다<sup>17)</sup>. 그러나 1차 실험은 1회성 조사였기 때문에, 패널 자료가 지닌 신뢰성 문제의 한 축인 표본 이탈 부분을 검증할 수가 없었다. 홍민기 외(2014)<sup>18)</sup>에 의하면, 패널의 표본이탈은 이탈비율이 높지 않더라도 조사 기간 동안 누적적으로 발생함으로써 자료의 신뢰성 및 대표성에 영향을 미칠 수 있기 때문에 중요하다고 언급하였다.

노동패널탐에서 웹조사 실험 연구를 진행하는 목적은 현재 사용 중인 CAPI를 CAWI로 전면 교체하기 위함이 아니다. 기존의 CAPI를 계속 유지하면서, 대면조사의 한계를 극복하기 위한 대안으로서 CAWI를 검토 중이다. 따라서 조사도구별로 표본이탈이 다르게 나타나 유지집단과 탈락집단의 성격이 조사도구별로 다르다면, 즉, 자료의 신뢰성 및 대표성에 영향을 미친다면, CAWI의 도입도 재검토되어야만 한다.

이에 동일한 표본을 대상으로 2차 웹조사 실험을 기획하였다.

2차 CAWI 실험의 목적은 ① 조사도구별 유지율의 차이와 표본 이탈의 편향이 발생하는가? 즉 유지된 집단과 탈락한 집단이 조사도구별로 다른 특성을 가지고 있는가? ② 2차 웹조사 실험 결과는 1차 조사 결과와 어떠한 차이가 있는가? ③ 1차 실험 결과에 따라 수정된 설문문항의 경우, 2차 웹조사 결과와는 차이가 있는가?

이상의 목적에 따라 제2절에는 2차 CAWI 실험 설계에 대해 서술하고, 제3절에는 1차와 2차 조사 결과를 실증적으로 분석, 제4절에는 1차 보고

17) 1차 웹조사 실험연구 결과에 대한 자세한 내용은 제4절 소결을 참조하기 바란다.

18) 홍민기 외(2014), 『패널자료 품질개선 연구(III)』의 제5장 『Nonignorable panel attrition에 대한 보정 연구』(김재광 저술)에 언급된 내용이다.

서내용과 2차 보고서내용을 정리하였다.

## 제2절 노동패널 CAWI 2차 실험 설계 개요

노동패널 CAWI 2차 실험은 다음의 네 단계로 이루어졌다.

첫째, 조사 설문지를 확정하였다. 2차 실험 조사 설문지는 1차 조사와 동일하게 구성하였다. 기존의 노동패널 본조사는 1차조사 설문지와 2차조사 설문지가 다르게 구성되어 있다. 1차에는 가구 설문과 신규조사자용인 개인 설문으로 구성된다. 2차 이후부터는 가구 설문은 동일하고, 개인 설문은 응답자 유형에 따라 달라진다. 지난 조사 이후 일자리 종류와 일자리 지속 유무를 기준으로 유형 설문(유형 1~8)과 취업자 설문, 미취업자 설문, 총 10가지 종류의 설문 중에서 선별하여 질문한다. 즉 1차 조사에서 조사된 일자리 유형에 따라 조사해야 할 설문의 종류가 달라지는 것이다. 노동패널 개인조사가 복잡한 설문구조를 가지고 있는 이유는 개인의 일자리를 추적 조사하여 그 특성을 파악하기 위한 연구목적에 의해 설계되었기 때문이다. 그러나 2차 웹조사 실험의 목적은 개인의 일자리 추적이 아니라 1차 실험조사와 2차 실험조사의 결과를 비교 분석하여 조사도구별 차이를 파악하고, CAWI가 대면조사인 CAPI의 한계를 극복할 수 있는 대안으로 가능한 것인가를 검토하는 것이다. 만약 2차 웹조사 실험에서 노동패널 본조사와 동일하게 개인 설문을 유형 설문으로 진행하면, 비교대상인 1차 조사와 설문문항이 달라져서 각 차수의 조사 결과를 비교 검증하는 데 한계가 발생할 수밖에 없다. 따라서 2차 웹조사 실험 연구에서는 1차조사와 동일하게 지난 조사 여부와 관계없이 모두 신규조사자용 개인 설문을 사용하여 1차 조사와 2차 조사의 비교 분석이 가능하도록 하였다.

둘째, 1차 CAWI 실험 조사 프로그램의 오류를 수정 보완하고, 추가적인 실험을 위해 프로그램을 업데이트하였다. 1차 실험에서는 기존 노동패널 본조사에서 사용하는 BLAISE 기반 프로그램이 아니라 한국리서치에서 개발한 프로그램을 사용하였다. 그 이유는 한국의 많은 조사업체가 자



체 프로그램들을 개발하여 사용하고 있는 만큼, 노동패널에서 20년 전에 개발된 외국 프로그램을 굳이 고집할 이유가 없기 때문이다. 그러나 프로그램을 수정할 때 발생할 수 있는 오류를 차단하기 위해 BLAISE를 계속 사용하고 있는바, 새롭게 개발된 한국리서치 프로그램이 오류 없이 작동할 수 있는지 내부적 검토가 필요하였다. 1차 실험 조사에서 새로운 프로그램들은 무리 없이 작동하였고, 다만 혼인 문항과 관련된 로직 오류가 발견되어 CAWI 프로그램 오류를 수정 보완하였다.

추가적인 실험을 위한 지시문 변경은 사회보험 항목, 산업 및 직업, 근로기간과 관련된 문항이다. 구체적인 변경 내용은 개별 분석에서 다시 다룰 것이다.

셋째, 조사 규모를 확정하고 조사를 시행하였다. 2차 실험 조사의 조사 대상은 1차에 조사된 CAPI 151가구와 CAWI 307가구와 그 가구원 전수이다. 거절 가구에 대해서는 1차 조사의 지역별 가구원수별 할당을 유지하면서, 한국리서치 MS(Master Sample) 패널 대상으로 추가 표집하였다. 전체 응답가구는 539가구였고, 이 중 분석에 사용된 1차 2차에 모두 응답한 균형패널 가구는 303가구이다<sup>19)</sup>.

조사는 노동패널팀이 한국리서치에 의뢰하여 진행되었고, 조사기간은 2017년 12월~2018년 1월이다.

넷째, 1차와 2차 실험 조사 결과를 분석하였다. 1차 조사는 조사도구별 집단의 차이를 검증하는 것을 1차 목표로 하였다. 2차 조사 결과는 1차 조사 결과와 함께 비교 검증하여 조사도구별, 차수별 차이가 발생하는지를 알아보았다. 주요 분석 항목은 인구학적 특성에 따른 유지율 차이, 항목응답률, 분포의 차이, 응답값의 차이, 설문 유형에 따른 차이, 직산업 분류 차이 등이다.

세부 항목으로는 가구의 소득이나 자산, 생활비 등 가구 변수와 일자리

19) 조사도구별 2차 응답가구 수는 아래와 같다.

	기존	신규	합계
CAPI	95	67	162
CAWI	208	169	377
합계	303	236	539

정보와 근로소득, 만족도, 술, 담배 등 건강 관련 변수, 사회보험 수급 여부 등 개인 변수들을 사용하였다.

분석방법은 연속변수와 항목응답률, 항목무응답률 등을 검정할 때에는 t-검정(t-test)을 사용하였다. 분할표(contingency table)의 경우에는 피셔의 정확검정(Fisher's exact test) 또는 독립성검정( $\chi^2$  검정)을 사용하였다. 또한 분산의 차이가 있는지를 확인할 때에는 분포형태와 관계없이 검정할 수 있는 Levene 검정을 실시하였다.

### 제3절 CAWI 1차 2차 조사에 관한 실증분석

#### 1. 표본유지율

##### 가. 표본의 유지 및 자료 오류

2차 실험 조사 결과는 <표 3-1>과 같다. 1차 성공 CAPI 151가구 중 95가구(62.9%)가 응답하였고, CAWI는 307가구 중 208가구(67.8%)가 응답하여 CAWI의 유지율이 약간 높았다. 개인은 CAPI 255명 중 140명(54.9%), CAWI 509명 중 279명(54.8%)이 응답하여 거의 비슷하였다. 가구와 개인 모두 유지율에 있어서는 조사도구별 차이가 통계적으로 유의하지 않아, CAPI와 CAWI의 혼합 사용이 패널 유지에는 차이가 없을 것으로 생각된다.

신뢰성 있는 노동패널 자료를 만들기 위해서 많은 사람들이 복잡한 절차를 거쳐 노력한다. 노동패널 본 조사 자료 생성과정을 살펴보면, 가구에 방문하기 전, 전년도 응답내용 중 중요한 부분을 미리 파악할 수 있는 인포시트(Information sheet의 줄임말)를 출력한다. 면접원이 조사 전에 인포시트를 확인하면, 가구원의 이름과 나이, 일자리 정보, 가구원번호 등을 알 수 있어, 조사 시 발생할 수 있는 오류를 줄일 수 있다. 설문이 완료되면, 1차적으로 각 지방 사무소에 배치되어 있는 리뷰어(reviewer)가 IRP(Intelligent Review Program)를 출력해 전년도 응답내용과 올해 응

〈표 3-1〉 조사도구별 패널유지율 차이

		1차 성공	2차 유지	조사 유지율	S.D	diff <sup>(1)</sup>	t-value	Pr> t
가구	CAPI	151	95	0.6291	0.4846	-0.0484	-1.0277	0.3046
	CAWI	307	208	0.6775	0.4682			
개인	CAPI	255	140	0.5490	0.4986	0.0009	0.0232	0.9815
	CAWI	509	279	0.5481	0.4982			

주: 1) diff는 CAPI 조사시작률에서 CAWI 조사시작률을 뺀 차이를 의미함.

답내용을 비교한다. 이때 이상치가 발생하면, 면접원에게 다시 응답자와 연락하여 확인하도록 한다. 확인이 끝난 자료는 한국리서치 연구부서에 서 독자적으로 개발한 KLIPS Clean(노동패널 데이터 클리닝 프로그램)을 사용하여 2차 검증을 실시한다. 이때 발견된 오류도 다시 면접원에게 전달되어 응답자의 재확인을 받는다. 이렇게 원자료가 생성되고, 이를 전달받은 한국노동연구원 노동패널팀에서 약 1년 동안 수차례에 걸친 횡단·종단 정제작업을 실시하며, 한국리서치와 협력하여 오류를 수정하고, 이용자가 편하게 사용할 수 있도록 가공 자료를 생성한다.

그러나 본 실험 연구 자료는 앞서 언급한 노동패널만의 엄격한 자료 정제과정(인포시트 출력, IRP 검토, KLIPS Clean 프로그램)을 거치지 않고, 기본적인 확인 후 노동패널팀에 전달되었다. 따라서 정제 전에 발생된 오류들을 확인할 수 있으며, 오류의 양과 질에 대해 조사도구별로 차이가 나는지 검토할 수 있다. 세부 문항별 오류는 해당 분야에서 설명하고, 본 절에서는 패널 자료의 기본인 개인 식별 오류의 차이만 검증하였다. 오류율은 CAPI 341명 중 42명(12.3%)이고, CAWI 599명 중 55명(9.8%)이었으나, 통계적으로 유의하지는 않았다.

1차 자료와 2차 자료의 개인을 연결시키는 과정에서 발생한 개인 식별 오류 유형을 살펴보았다(그림 3-1 참조). 동일한 개인식별번호(pid)가 1차와 2차에 다른 사람에게 부여된 경우를 오류로 보고 정제하였다. 개인을 식별할 때는 이름, 성별, 연령, 생년월일을 주로 사용하여 확인하였고, 학력은 보조 변수로 사용하였다.

오류 유형 중 (1) 가구원 내 가구원번호의 오류란, A라는 사람이 1차

〈표 3-2〉 2차 조사의 개인 식별 변수(pid) 매칭 오류

	사례수	오류건수	오류율	S.D	diff <sup>1)</sup>	t-value	Pr< t
CAPI	341	42	0.1232	0.3291	0.0184	1.1624	0.2454
CAWI	559	55	0.0983	0.2981			

주: 1) diff는 CAPI 조사시작률에서 CAWI 조사시작률을 뺀 차이를 의미함.

〈그림 3-1〉 개인 식별 변수(pid) 매칭 오류 유형

오류 유형	1차		2차		2차 수정 내용	
	이름	가구원 번호	이름	가구원 번호	이름	가구원 번호
(1) 가구원 내의 가구원번호 오류	홍길동	1	홍길순	1	홍길동	1
	홍길순	2	홍길동	2	홍길순	2
(2) 가구원 추가 시 가구원번호 오류	홍길동	1	홍길동	1	홍길동	1
	홍길순	2				
			홍길자	2	홍길자	3
(3) 이름 오류	홍길동	1	홍길도	1	홍길동	1
(4) 가구 전체 모두 다른 이름	홍길동	1	이철수	1	수정불가	
	홍길순	2	이영희	2	수정불가	

〈표 3-3〉 2차 조사의 개인 식별 변수(pid) 매칭 오류 유형

오류 유형	CAPI	CAWI	전 체	Pr<=P
(1) 가구원 내의 가구원번호 오류	28( 66.7)	20( 36.4)	48( 49.5)	0.002
(2) 가구원 추가 시 가구원번호 오류	5( 11.9)	11( 20.0)	16( 16.5)	
(3) 이름 오류	8( 19.0)	8( 14.5)	16( 16.5)	
(4) 가구 전체 모두 다른 이름	0 -	9( 16.4)	9( 9.3)	
(5) 기타	1( 2.4)	7( 12.7)	8( 8.2)	
전 체	42(100.0)	55(100.0)	97(100.0)	

주: Fisher's exact 분석을 사용함.

조사에서 1번 가구원으로 응답하였는데, 2차 조사에서는 2번 가구원으로 응답하는 경우이다. 이는 앞서 노동패널 본 조사의 조사 과정 설명에서 언급한 인포시트만 제공되었다면 제대로 매칭되어 발생하지 않을 오류이다. 해당 오류가 전체 오류의 거의 절반을 차지하였고, CAPI는 66.7%, CAWI는 36.4%였다. (2) 가구원 추가 시 가구원번호 오류는 다음과 같다. A라는 사람이 1차에 1번 가구원으로 응답하였으나 2차에는 응답하지 않고 1

차에는 응답하지 않았던 신규 가구원이 2차에 응답하는 경우, 또 가구원 번호를 3번을 부여받아야 함에도 2번을 부여받는 경우이다. 1차 자료와 2차 자료를 결합하는 과정에서 A와 B가 모두 2번 가구원으로 매칭되어 발생한 오류이다. 이 오류도 인포시트가 있었다면, 가구원번호를 확인하고 정확한 가구원번호를 부여할 수 있어 발생빈도는 현저히 낮아질 것이라 생각된다. (3) 이름 오류는 한 음절에서 오타가 난 것으로 보인다. (4) 가구 전체 모두 다른 이름이라는 것은 한 가구의 가구원 모두가 바뀐 경우이다. 이 경우는 CAWI에서만 발생했는데, 원가구원이 없는 상태에서 모두 신규 가구원으로 조사되었기 때문에 자료를 사용할 수 없고, 가구 자료와 개인 자료 모두 분석에서 제외하였다. 다른 오류 유형은 인포시트와 같이 조사과정상에서 보완될 수 있는 것과 달리 (4)번 오류는 자료의 신뢰성을 떨어뜨리는 심각한 오류라고 생각된다.

오류 유형은 CAPI와 CAWI가 통계적으로 차이가 있다고 분석되었다.

#### 나. 가구원수, 지역에 따른 가구유지율 차이

일차적으로 패널조사는 설문응답자가 이탈하지 않고, 계속적으로 응답하는 것이 중요하다. 그러나 자료의 신뢰성을 보장하기 위해 패널유지율보다 더 중요한 것은 계속 응답하는 패널과 탈락하는 패널 간의 집단 간 차이가 발생하지 않아야 한다. 이에 조사도구별 패널유지집단과 탈락집단 간 차이를 가구 특성에 따라 살펴보고자 한다.

CAPI 유지집단은 CAPI로 응답된 가구들 중 1차, 2차 모두 응답한 집단으로 95가구이고, CAPI 탈락집단은 1차에는 응답했지만 2차에는 응답하지 않은 집단으로 56가구이다. CAWI도 CAPI와 마찬가지로, 1차 2차 응답 여부로 집단을 구분하여 유지집단 208가구, 탈락집단 99가구를 대상으로 분석하였다. 집단별로 보면, CAPI 탈락집단이 3.0명으로 가장 높았고, CAPI 유지집단이 2.7명으로 가장 적었으나, 통계적으로 유의하지는 않았다.

<표 3-5>와 <표 3-6>은 지역별, 집단별로 차이가 있는지 검증한 내용이다. 지역을 서울권, 경인권, 충청권, 영남권, 호남권으로 나누었을 때는

통계적으로 유의한 차이가 발생하였다. CAPI 유지집단에서는 서울권, 영남권, 경인권 순으로 많았고, CAPI 탈락집단은 경인권, 영남권, 서울권 순으로 많았다. CAWI 유지집단은 CAPI 유지집단과 달랐는데, 경인권, 영남권,

〈표 3-4〉 조사도구별 패널 유지 탈락별 가구원수 차이 검증

		사례수	가구원수	S.D	F-value	Prob>F
CAPI	유지	95	2.663158	1.234353	1.22	0.3022
	탈락	56	3.035714	1.36134		
CAWI	유지	208	2.730769	1.245132		
	탈락	99	2.838384	1.25118		

〈표 3-5〉 조사도구별 패널 유지 탈락별 지역 차이 1

		서울권	경인권	충청권	영남권	호남권	Pearson chi <sup>2</sup>	Pr
CAPI	유지	33 (34.7)	15 (15.8)	9 ( 9.5)	32 (33.7)	6 ( 6.3)	22.9164	0.028
	탈락	14 (25.0)	18 (32.1)	6 (10.7)	16 (28.6)	2 ( 3.6)		
CAWI	유지	39 (18.8)	67 (32.2)	19 ( 9.1)	55 (26.4)	28 (13.5)		
	탈락	27 (27.3)	32 (32.3)	8 ( 8.1)	24 (24.2)	8 ( 8.1)		

〈표 3-6〉 조사도구별 패널 유지 탈락별 지역 차이 2

		수도권	광역시	도	Pearson chi <sup>2</sup>	Pr
CAPI	유지	47 (49.5)	23 (24.2)	25 (26.3)	7.5374	0.274
	탈락	27 (48.2)	13 (23.2)	16 (28.6)		
CAWI	유지	98 (47.1)	46 (22.1)	64 (30.8)		
	탈락	56 (56.6)	27 (27.3)	16 (16.2)		

서울권 순이었고, CAPI 탈락집단은 경인권, 서울권, 영남권 순이었다.

지역을 수도권, 광역시, 도로 구분하였을때는 통계적 차이가 발견되지 않았다.

#### 다. 성별, 연령, 학력, 동거 여부에 따른 개인유지율 차이

CAPI 유지집단은 CAPI로 응답된 개인들 중 1차, 2차 모두 응답한 집단이 140명이고, CAPI 탈락집단은 1차에 응답했지만 2차에는 응답하지 않은 집단으로 115명이다. CAWI도 CAPI와 마찬가지로, 1차 2차 응답여부로 집단을 구분하여 유지집단 279명, 탈락집단 230명을 대상으로 분석하였다.

조사도구별 성별 유지집단과 탈락집단을 살펴보면, 4개 집단 모두 남성이 여성보다 많이 응답하였고, 특히 CAWI 유지집단에서 남성의 비율이 높았다. 그러나 4개 집단에 대한 독립성검정을 실시한 결과 통계적으로 유의하지는 않았다.

연령별로 살펴보면, 조사도구와 상관없이, 유지집단의 연령이 더 높은 것을 알 수 있다. 특히 CAPI 유지집단이 44.8세로 가장 높았고, CAWI 탈락집단이 39.4세로 가장 낮았다. 패널조사에 CAWI 조사방법을 고려하기 시작한 이유 중 하나가 해를 거듭할수록, 젊은 연령층의 이탈이 상대적으로 많았기 때문이다. 이와 같은 사실을 감안하면, CAWI 탈락집단의 평균 연령이 가장 낮은 것은 CAWI 도입을 통해 젊은층의 탈락률을 낮추는 데 효과가 크지 않을 것이라 생각이 든다.

집단별 학력을 살펴보면, 통계적으로 유의한 차이가 발생하지 않았다.

집단별 동거 여부는 각 셀의 사례수가 적어 해석에 유의해서 살펴보면, 통계적으로 유의하지 않았다.

〈표 3-7〉 조사도구별 패널 유지 탈락별 성별 차이

		여성	남성	Pearson $\chi^2$	Pr
CAPI	유지	66 (47.1)	74 (52.9)	4.6576	0.199
	탈락	56 (48.7)	59 (51.3)		
CAWI	유지	113 (40.5)	166 (59.5)		
	탈락	113 (49.1)	117 (50.9)		

〈표 3-8〉 조사도구별 패널 유지 탈락별 연령 차이

		사례수	연령	S.D	F-value	Prob>F
CAPI	유지	140	44.80714	12.90914	5.36	0.0012
	탈락	115	41.38261	13.2801		
CAWI	유지	279	42.07885	11.98264		
	탈락	229	39.36681	13.54479		

〈표 3-9〉 조사도구별 패널 유지 탈락별 학력 차이

		중졸	고졸	전문대졸	대졸	석사 이상	Pearson chi <sup>2</sup>	Pr
CAPI	유지	10 ( 7.1)	21 (15.0)	17 (12.1)	71 (50.7)	21 (15.0)	14.8047	0.252
	탈락	9 ( 7.8)	20 (17.4)	22 (19.1)	48 (41.7)	16 (13.9)		
CAWI	유지	24 ( 8.6)	52 (18.6)	52 (18.6)	125 (44.8)	26 ( 9.3)		
	탈락	27 (11.7)	29 (12.6)	49 (21.3)	102 (44.3)	23 (10.0)		

〈표 3-10〉 조사도구별 패널 유지 탈락별 동거 여부 차이

		비동거	동거	Fisher's exact
CAPI	유지	1 (0.7)	139 (99.3)	0.293
	탈락	3 (2.6)	112 (97.4)	
CAWI	유지	10 (3.6)	269 (96.4)	
	탈락	9 (3.9)	221 (96.1)	

## 2. 가구 및 개인의 특성 차이

본 장에서는 1차년도와 2차년도를 모두 응답한 가구만을 대상으로 분석하였다. 이는 주요 목적인 조사도구 간 비교뿐만 아니라 조사차수 간 비교도 살펴볼 수 있다. 조사도구 간 비중은 CAPI의 경우 31.4%(95가구)이고, CAWI의 경우 CAPI보다 2.2배 정도 많은 68.7%(208가구)로 나타났다.

### 가. 가구소득변수

2차년도 가구 설문은 1차년도와 기본적으로 동일한 문항으로 구성되어



있다. 먼저 가구소득의 항목은 총 6가지이며, 항목은 근로소득, 금융소득, 부동산소득, 사회보험소득, 이전소득 및 기타소득이다. 가구소득은 작년 한 해 동안과 지난 한 달 동안을 기준으로 하여 질문한다. 근로소득을 제외한 5가지의 나머지 소득 문항들은 서브문항(subquestion)으로 세분화하여 조사하고 있다. 본 보고서에서는 여섯 가지 소득별로 항목응답률이 CAPI와 CAWI 간 차이가 있는지 조사 차수별로 분석하였다. 참고로 2차년도에서의 사회보험소득액도 1차년도 조사 결과처럼 조사도구와 상관없이 수혜받고 있는 가구가 없어 분석에서 제외하였다.

<표 3-11>을 보면 1차년도 작년 한 해 근로소득액의 응답률이 CAPI가 91.6%이고 CAWI는 84.6%였다. 지난 한 달 근로소득액 역시 CAPI가 CAWI보다 4.2% 정도 높은 것으로 나타나 가구 근로소득액의 응답률은 모두 CAPI가 높은 편으로 나타났다. 2차년도의 작년 한 해 근로소득 및 지난 한 달 근로소득에 대한 응답률은 1차년도와 다르지만 비슷한 결과를 가지며, 조사도구와 상관없이 모두 1차년도 응답률보다 높은 것으로 나타났다.

가구소득 응답률이 조사도구 간 차이가 있는지를 분석한 결과인 <표 3-12>를 보면, 1차년도의 경우 금융소득액을 제외하면 모두 CAPI에서 응답률이 높게 나타났다. 부동산소득액 및 이전소득액의 응답률은 조사도구 간 통계적으로 유의한 차이가 있는 것으로 분석됐다. 또한, 근로소득액도 유의수준 10%에서는 통계적으로 유의한 차이가 있어서 금융소득액과 기타소득액을 제외하고는 응답률이 조사도구 간 차이가 있다고 볼 수 있다. 근로소득액, 금융소득액, 부동산소득액, 이전소득액과 기타소득액을 모두 합한 값인 생성변수 소득총액은 CAPI의 응답률이 CAWI보다 4.24%포인트 높은 98.95%이고, t-test 결과 통계적으로 유의한 차이가 있

<표 3-11> 가구 근로소득액 응답률(CAPI vs CAWI)

		CAPI	CAWI
1차	작년 한 해 근로소득	91.6	84.6
	지난 한 달 근로소득	87.4	83.2
2차	작년 한 해 근로소득	93.7	88.9
	지난 한 달 근로소득	92.6	84.6

는 것으로 나타났다.

2차년도의 경우 기타소득액을 제외하고는 모두 CAPI에서 응답률이 높게 나타났다. 부동산소득액, 이전소득액(유의수준 10%)과 소득총액(유의수준 10%)은 1차년도 결과와 동일하게 통계적으로 유의한 차이가 있는 것으로 나타났다.

2차년도의 근로소득액과 금융소득액(CAPI만 해당)을 제외한 나머지 항목에서는 조사도구별 및 항목별 1차년도 응답률보다 모두 낮은 것으로 나타났다. 특히 이전소득의 경우 CAPI에서는 1차년도에 50.53%였으나 38.95%포인트 하락하여 2차년도에 11.58%였으며, CAWI에서는 1차년도에 23.08%에서 17.79%포인트 하락한 5.29%(2차년도)로 큰 폭의 감소세를 보였다.

〈표 3-12〉 가구소득 응답률의 집단(CAPI vs CAWI) 간 차이

		CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t
		평균	S.D	평균	S.D			
1차	소득총액 <sup>2)</sup>	0.9895	0.1026	0.9471	0.2243	0.0424	2.26	0.0248
	근로소득액	0.9158	0.2792	0.8462	0.3617	0.0696	1.83	0.0687
	금융소득액	0.4316	0.4979	0.4615	0.4997	-0.03	-0.48	0.6282
	부동산소득액	0.1895	0.394	0.0865	0.2818	0.1029	2.29	0.0234
	이전소득액	0.5053	0.5026	0.2308	0.4223	0.2745	4.63	<.0001
	기타소득액	0.3368	0.4751	0.3077	0.4627	0.0291	0.5	0.6143
2차	소득총액 <sup>2)</sup>	0.9789	0.1443	0.9423	0.2337	0.0366	1.67	0.0962
	근로소득액	0.9368	0.2445	0.8894	0.3144	0.0474	1.43	0.155
	금융소득액	0.4737	0.502	0.4567	0.4993	0.017	0.27	0.7845
	부동산소득액	0.1684	0.3762	0.0721	0.2593	0.0963	2.26	0.0253
	이전소득액	0.1158	0.3217	0.0529	0.2243	0.0629	1.72	0.0869
	기타소득액	0.2526	0.4368	0.2837	0.4519	-0.031	-0.57	0.5712

주: 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

2) 소득총액은 근로소득액, 금융소득액, 부동산소득액, 이전소득액과 기타소득액을 모두 합한 값임.

## 나. 금융자산 변수

금융자산은 6개의 서브문항을 가지며, 이를 하나의 변수로 통합하여 분석하였다. 조사차수별 응답률의 집단(CAPI vs CAWI) 간 차이를 분석한 결과, 1차년도에 응답률은 조사도구 간 통계적으로 유의한 차이가 있는 것으로 나타났다. CAPI의 응답률이 CAWI보다 10.78%포인트 높다. 2차년도에는 조사도구와 상관없이 모두 응답한 것으로 나타났다.

〈표 3-13〉 금융자산 응답률의 집단(CAPI vs CAWI) 간 차이

	CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t
	평균	S.D	평균	S.D			
1차	0.9684	0.1758	0.8606	0.3472	0.1078	3.58	0.0004
2차	1.0	0	1.0	0	0	-	-

주: 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

## 다. 지난 한 달 소득변수

지난 한 달 소득의 서브문항은 6개를 가지며, 금융자산과 같이 하나의 변수로 통합하여 분석하였다. 1차년도에서 응답률 집단(CAPI vs CAWI) 간 차이를 분석한 결과, 응답률은 집단 간 통계적으로 유의한 차이가 있는 것으로 나타났으며 CAPI가 CAWI보다 응답률이 9.43%포인트 높다. 2차년도 역시 동일한 결과를 가진다.

〈표 3-14〉 지난 한 달 소득 응답률의 집단(CAPI vs CAWI) 간 차이

	CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t
	평균	S.D	평균	S.D			
1차	0.9789	0.1443	0.8846	0.3203	0.0943	3.53	0.0005
2차	0.9684	0.1758	0.8846	0.3203	0.0838	2.93	0.0037

주: 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

## 라. 월평균 저축액

월평균 저축액은 2가지 항목(저축액 전체를 묻는 항목, 8개의 서브문항을 묻는 항목)으로 구성된다. 8개의 서브문항을 모두 합한 값은 저축액 전체와 동일하므로, 저축액 전체를 묻는 항목에 대해 분석하였다. 1차년도에서 응답률의 집단(CAPI vs CAWI) 간 차이를 분석한 결과는 유의수준 10%에서 통계적으로 유의한 차이가 있으며, 응답률은 CAPI가 CAWI보다 7.83%포인트 높은 것으로 나타났다. 2차년도 역시 조사도구 간 응답률의 차이가 통계적으로 유의하였다.

〈표 3-15〉 월평균 저축액 응답률의 집단(CAPI vs CAWI) 간 차이

	CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t
	평균	S.D	평균	S.D			
1차	0.9053	0.2944	0.8269	0.3792	0.0783	1.96	0.0516
2차	0.9053	0.2944	0.8029	0.3988	0.1024	2.5	0.0131

주: 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

## 마. 생활비 및 주거 관련 변수

가구 생활비 및 주거 관련 변수에 대해 살펴보았다. 먼저, 가구 생활비는 작년 한 해 월평균 생활비 전체를 묻는 간단한 문항과 20개의 서브문항을 묻는 자세한 문항으로 구성되어 있다. 주거 관련 변수는 입주 형태 및 주택 종류에 따라 각각 구분하여 대지면적, 평수, 시가 임대보증금과 월세금 등과 관련하여 질문하고 있다.

가구 생활비의 경우 전체 및 생활비 합계(생성변수)를 모두 살펴보았고, 주거 관련 변수의 경우 주거형태를 2개의 집단(자가, 전월세)으로 구분하여 분석하였다.

가구 생활비를 살펴보면 1차년도를 통틀어 응답률의 집단 간 차이 면에서 통계적으로 유의한 차이가 없는 것으로 나타났다. 생활비 합계의 경우 조사도구와 상관없이 응답률은 모두 100%였다.

2차년도 전체의 경우 1차년도와 동일한 결과를 가지며, 생활비 합계의 경우 유의수준 10%에서 조사도구 간 응답률 차이가 통계적으로 유의한 것으로 나타났다.

<표 3-17>의 주거지 정보에 대한 응답률의 집단 간 차이는 1차년도 및 2차년도에서 모두 통계적으로 유의한 차이가 없는 것으로 나타났다.

<표 3-16> 가구 생활비 응답률의 집단(CAPI vs CAWI) 간 차이

		CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t
		평균	S.D	평균	S.D			
1차	전 체	1.0	0	0.9952	0.0693	0.00481	1	0.3185
	생활비 합계	1.0	0	1.0	1.0	0	-	-
2차	전 체	0.9895	0.1026	0.9856	0.1195	0.0039	0.27	0.7836
	생활비 합계	1.0	0	0.9856	0.1195	0.0144	1.74	0.0833

주: 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

<표 3-17> 주거지 정보 응답률의 집단(CAPI vs CAWI) 간 차이

		CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t
		평균	S.D	평균	S.D			
1차	자가: 평수	0.5368	0.5013	0.5337	0.5001	0.0032	0.05	0.959
	자가: 평당 시가	0.6105	0.4902	0.5817	0.4945	0.0288	0.47	0.6376
	전월세: 평수	0.3895	0.4902	0.4183	0.4945	-0.0288	-0.47	0.6376
	전월세: 임대보증금	0.3895	0.4902	0.4183	0.4945	-0.0288	-0.47	0.6367
	전월세: 월세	0.3895	0.4902	0.4183	0.4945	-0.0288	-0.47	0.6376
2차	자가: 평수	0.5579	0.4993	0.5769	0.4952	-0.019	-0.31	0.7572
	자가: 평당 시가	0.6421	0.4819	0.625	0.4853	0.0171	0.29	0.7756
	전월세: 평수	0.3579	0.4819	0.375	0.4853	-0.0171	-0.29	0.7756
	전월세: 임대보증금	0.3579	0.4819	0.375	0.4853	-0.0171	-0.29	0.7756
	전월세: 월세	0.3579	0.4819	0.375	0.4853	-0.0171	-0.29	0.7756

주: 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

2) 자가에서 단독주택의 경우는 평수 대신 대지면적과 연건평을 질문하여 같이 분석할 수 없으므로 제외함.

### 바. 일자리 관련 변수

1차와 2차 조사에서 일자리 관련 변수 중 임금근로자의 비중과 종업원 수 응답률의 조사방법에 따른 차이를 제시했다. <표 3-21>을 보면 1차와 2차 모두 CAPI에 비해 CAWI에서 임금근로자 비중이 높게 나타났으나 모두 통계적으로 유의하지는 않았다.

<표 3-18>은 차수별 조사방법에 따른 종업원수 응답률 차이를 비교한 것이다. 임금근로자의 조사방법에 따른 응답률 차이는 약 1% 내외로 나타났고 통계적으로도 유의하지 않았다. 한편 비임금근로자는 CAPI에서 1, 2차 모두 종업원수 응답률이 100%로 나타났고 CAWI에서는 그보다 낮게 나타났다. 다만 비임금근로자 표본이 적어 통계적으로는 조사방법에 따른 차이가 없는 것으로 나타나고 있다.

<표 3-18> 임금근로자 비중 조사방법에 따른 차이 비교

	CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t
	임금근로자 비중	S.D.	임금근로자 비중	S.D.			
1차	0.8039	0.3990	0.8544	0.3536	-0.0504	1.13	0.2600
2차	0.8182	0.3877	0.8424	0.3653	-0.0242	0.53	0.5970

주: 1) diff는 각 차수별 CAPI와 CAWI의 임금근로자 비중 차이를 의미.

<표 3-19> 종업원수 응답률의 조사방법에 따른 차이 비교

		CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t
		응답률	S.D.	응답률	S.D.			
1차	임금근로자	0.7927	0.4079	0.8011	0.4003	-0.0085	0.16	0.8754
	비임금근로자	1.0000	-	0.9333	0.2537	0.6667	1.17	0.2473
2차	임금근로자	0.8148	0.3909	0.8012	0.4003	0.0136	0.25	0.7992
	비임금근로자	1.0000	-	0.9688	0.1768	0.0313	0.75	0.4589

주: 1) diff는 각 차수별 CAPI와 CAWI의 응답률 차이를 의미.

## 사. 흡연 및 음주

한국노동패널조사의 설문 문항 중에는 개인의 성향이나 개인의 의견 등과 같이 응답자에게 민감한 질문도 포함되어 있다. 이에 개인의 성향을 묻는 흡연 및 음주에 대한 응답비중으로 집단 간 차이가 있는지 살펴보았다. 우선 <표 3-20>에서는 흡연 및 음주 항목에서 무응답이 관측되지 않았지만 응답률(‘있음’에 대한 응답비중)의 CAPI와 CAWI 집단 간 차이를 분석한 결과이다. 1차년도 조사에서는 5% 수준에서 집단 간 통계적으로 유의한 차이가 있는 변수는 흡연 및 금연 변수이고, 2차년도 조사에서도 1차년도와 마찬가지로 흡연 및 금연 항목은 5% 수준에서 유의한 차이가 있는 것으로 보이며, 1차년도와 상이한 금주 변수는 통계적으로 유의한 차이가 있는 것으로 나타났다.

<표 3-20> 흡연 및 음주 응답비중의 집단 간 차이

		CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t
		평균	S.D.	평균	S.D.			
1차	흡연	0.1000	0.3011	0.2043	0.4039	-0.1043	-2.97	0.0032
	금연	0.0286	0.1672	0.0932	0.2912	-0.0646	-2.88	0.0042
	음주	0.6643	0.4739	0.6595	0.4747	0.0048	0.10	0.9224
	금주	0.0214	0.1453	0.0323	0.177	-0.0108	-0.67	0.5049
2차	흡연	0.0929	0.2913	0.2222	0.4165	-0.1294	-3.69	0.0003
	금연	0.0357	0.1862	0.0968	0.2962	-0.0611	-2.58	0.0104
	음주	0.6429	0.4809	0.6703	0.471	-0.0274	-0.56	0.5774
	금주	0.0071	0.0845	0.0609	0.2396	-0.0538	-3.36	0.0009

주: 1) diff는 CAPI와 CAWI의 응답률 차이를 의미.

2) 1차, 2차년도 균형패널 사용.

## 아. 직업훈련, 사회보험 및 혼인상태 변화

<표 3-21>는 응답의 번거로움이나 사회적인 민감성으로 인해 응답기 피가 발생할 수 있는 개인적 민감 문항인 직업훈련 경험 여부, 자격증 보

유 여부, 사회보험 수급 여부, 혼인상태 변화 여부 항목에 대한 응답률이 다. 4개 항목에서 무응답은 관측되지 않았으며 응답률(‘있음’에 대한 응답 비중)의 집단 간 차이를 분석한 결과, 1차년도 조사에서는 직업훈련 경험과 사회보험급여 수급 변수의 응답률이 집단 간 통계적으로 유의한 차이가 있는 것으로 보인다. 자격증 보유 항목은 통계적으로 유의하지 않은 것으로 나타났으며, 혼인상태 변화 여부 항목의 집단 간 응답률을 비교하였지만 통계적으로 차이가 없는 것으로 나타났다(김유빈 외, 2017).

2차년도 조사에서 통계적으로 차이가 있는 항목은 자격증 보유 여부 및 사회보험 수급 여부 항목이다. 추가로 직업훈련 경험이 있는 경우 직업훈련 받은 개수에 대한 응답률은 1차년도 및 2차년도 조사에서 통계적

〈표 3-21〉 민감 문항 응답률의 집단 간 차이

		CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t
		평균	S.D.	평균	S.D.			
1차	직업훈련 경험 여부	0.1255	0.3319	0.2554	0.4365	-0.1299	-4.57	<.0001
	직업훈련받은 개수	1.8500	0.9333	1.6447	0.8279	0.2053	0.96	0.3392
	자격증 보유 여부	0.4902	0.5009	0.4401	0.4969	0.0501	1.31	0.1902
	자격증 보유 개수	1.6267	1.1243	2.2016	1.5879	-0.5749	-3.01	0.0029
	사회보험 수급 여부	0.0980	0.2980	0.1807	0.3852	-0.0827	-3.27	0.0011
	수급사회보험 개수	1.0000	0.0000	1.4821	0.9907	-0.4821	-3.64	0.0006
	혼인상태 변화 여부	0.0627	0.2430	0.0806	0.2724	-0.0178	-0.92	0.3597
	혼인상태 변화 횟수	1.1000	0.3162	1.6522	0.7141	-0.5522	-3.08	0.0043
2차	직업훈련 경험 여부	0.1516	0.3593	0.2021	0.4020	-0.0505	-1.84	0.0665
	직업훈련받은 개수	1.7895	0.9177	1.7429	0.7743	0.0466	0.22	0.8236
	자격증 보유 여부	0.3755	0.4851	0.4615	0.4990	-0.0861	-2.37	0.0180
	자격증 보유 개수	1.7018	1.1490	2.2656	1.8076	-0.5639	-2.56	0.0115
	사회보험 수급 여부	0.1119	0.3158	0.1646	0.3711	-0.0527	-2.14	0.0329
	수급사회보험 개수	1.0000	0.0000	1.4000	0.8367	-0.4000	-3.21	0.0025
	혼인상태 변화 여부	0.0325	0.1776	0.0304	0.1719	0.00208	0.16	0.8707
	혼인상태 변화 횟수	1.7500	0.7071	1.5385	0.9047	0.2115	0.6	0.5497

주: 1) diff는 CAPI와 CAWI의 응답률 차이를 의미.

2) 1차, 2차년도 균형패널 사용.



으로 유의하지 않은 것으로 나타났지만, 자격증이 있는 경우 자격증 보유 개수에 대한 응답률은 CAPI와 CAWI 조사방식에 차이가 있는 것으로 보인다. 사회보험 수급을 받거나 받고 있는 경우에도 수급사회보험 개수에 대한 응답률 역시 통계적으로 차이가 있는 것으로 나타났다. 혼인상태 변화 여부나 변화 횟수의 응답률은 차이가 없는 것으로 보여진다.

### 3. 분포의 차이

#### 가. 입주형태, 주거형태 및 가구원수

<표 3-22>를 보면 1차년도와 2차년도의 경우 입주형태, 주택형태 및 가구원수는 유의수준 5%에서 조사도구 간 통계적으로 유의한 차이가 없음을 확인할 수 있었다. 유의수준 10%일 경우, 주택형태는 통계적으로 유의한 차이를 가진다. 2차년도와 관련한 검증 역시 1차년도 결과와 동일하게 나타났다.

<표 3-22> 입주·주거형태 및 가구원수 차이 검증

		CAPI: CAWI
1차	입주형태	0.6112
	주택형태	0.0637
	가구원수	0.4076
2차	입주형태	0.8687
	주택형태	0.0693
	가구원수	0.2609

주: 1) 분할표(contingency table) 독립성검정( $\chi^2$ )의 p값을 보고한 것임.

2) 입주형태는 자가/전월세/기타, 주택형태는 단독/아파트/연립·다세대/상가/기타, 가구원수는 1~6인임.

#### 나. 가구소득 및 금융자산

소득 및 자산 등과 같이 금액을 가지는 변수들의 평균에 대해 조사도구 간 차이가 있는지를 검증하였다. 금액 관련 변수는 값이 커서 입력해야 하는 숫자가 많아지거나, 또는 금액 단위에 대한 착각(예: 단위가 만

원이나 원을 기준으로 입력하여 응답값이 커지는 경우)으로 인하여 입력 오류가 발생할 가능성이 높다. 1·2차년도 조사를 살펴본 결과 이상치로 판단되는 경우가 있어서, 본 장에서는 이상치 여부를 고려하여 분석을 각각 실시하였다. 이때 이상치 판단 기준은 상한의 경우 99.5분위수 이상인 값으로, 하한의 경우 0.5분위수 이하인 값으로 정하였다.

먼저 이상치를 제외하지 않고 소득변수 차이를 검증한 결과인 <표 3-23>을 보면, 1·2차년도에서 모든 항목(근로소득액, 금융소득액, 부동산소득액, 이전소득액, 기타소득액)의 평균은 CAPI와 CAWI 간 통계적으로 유의한 차이가 없는 것으로 나타났다. 조사 차수와 상관없이 전반적으로 CAWI에서 소득변수들의 값이 더 넓게 퍼져 있어서 표준편차의 값도 크다는 것을 알 수 있다. 다음으로 분산에는 차이가 있는지를 확인하기 위해 분포형태와 관계없이 검정할 수 있는 Levene 검정도 실시하였다. 1·2차년도에서 모든 소득변수들의 분산은 CAPI와 CAWI 간에 통계적으로 차이가 없는 것으로 나타났다.

다음 <표 3-24>는 소득변수별 이상치 개수를 보여주고 있다. 1차년도

<표 3-23> 소득변수 차이 검증

		CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t	Pr> Z  <sup>2)</sup>	Levene 검정
		평균 (만 원)	S.D	평균 (만 원)	S.D					
1차	근로소득액	5,398.7	3,537.2	4,997.2	2,761.2	401.5	0.93	0.355	0.7345	0.1673
	금융소득액	329.2	617.0	1,501.0	7,528.6	-1,171.8	-1.51	0.1334	0.9456	0.496
	부동산소득액	1,270.6	1,218.3	6,907.8	20,324.5	-5,637.2	-1.17	0.2562	0.4286	0.2892
	이전소득액	2,231.9	5,880.9	7,951.6	43,243.4	-5,719.7	-0.91	0.3683	0.5427	0.3141
	기타소득액	1,376.7	4,599.3	7,004.5	45,600.8	-5,627.7	-0.98	0.332	0.3405	0.4771
2차	근로소득액	5,401.9	3,618.1	4,971.1	3,306.9	430.8	0.98	0.3284	0.4573	0.6727
	금융소득액	614.6	1,351.2	1,609.5	9,546.1	-994.9	-1	0.3221	0.9982	0.4871
	부동산소득액	3,111.3	4,361.2	3,601.3	7,301.2	-490.1	-0.23	0.8208	0.8586	0.428
	이전소득액	304.7	179.9	1,096.0	2,959.0	-791.3	-0.89	0.3966	0.531	0.2794
	기타소득액	4,548.4	15,989.9	1,260.1	5,698.5	3,288.3	0.99	0.3337	0.1045	0.1611

주: 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

2) Pr>|Z|는 비모수검정(nonparametric test)을 통해 계산된 값임.

〈표 3-24〉 소득변수에 대한 이상치 개수

(단위: 가구)

		CAPI	CAWI
1차	근로소득액	2	2
	금융소득액	0	2
	부동산소득액	1	1
	이전소득액	1	1
	기타소득액	0	2
2차	근로소득액	2	2
	금융소득액	1	2
	부동산소득액	1	1
	이전소득액	1	1
	기타소득액	1	1

금융소득액 및 기타소득액의 경우 CAWI는 이상치가 2개였으나, CAPI는 없었다. 2차년도의 경우 금융소득액을 제외하고는 CAPI와 CAWI의 이상치 개수가 동일하였다. CAWI에서는 주로 이상치가 상한값에서 발생한 반면에 CAPI에서는 하한값에서 발생하였다. 이는 CAWI의 경우 응답자가 직접 금액을 입력하는 과정에서 발생한 금액 단위에 대한 혼란이 하나의 요인이라고 볼 수 있다.

<표 3-25>는 이상치를 제외한 후 분석한 결과로, 1차년도 금융소득액 평균이 CAPI와 CAWI 간에 통계적으로 유의한 차이가 있는 것으로 나타났다. 나머지 소득변수들은 조사 차수와 무관하게 이상치 제거 전과 동일한 결과를 가진다.

이상치 제거 여부에 따른 소득변수의 평균 변화를 보면, CAWI의 경우 1·2차년도 금융소득액, 부동산소득액, 이전소득액에서는 이상치를 제거한 평균이 이상치를 제거하지 않은 평균보다 작으며 차이도 큰 편으로 나타났다. CAPI의 경우 2차년도 기타소득액을 제외한 나머지는 모두 평균 변화가 크지 않고 안정적이어서, 이상치의 영향이 CAWI보다 민감하지 않다고 볼 수 있다.

다음으로 차수별 소득변수의 평균 변화를 살펴보면, 소득변수 중에서

〈표 3-25〉 소득변수 차이 검증 : 이상치 제거 후 분석

		CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t	Pr> Z  <sup>2)</sup>	Levene 검정
		평균 (만 원)	S.D	평균 (만 원)	S.D					
1차	근로소득액	5,229.4	2,838.7	4,964.7	2,633.9	264.7	0.74	0.4598	0.7294	0.524
	금융소득액	329.2	617	756.4	1,531.2	-427.2	-2.31	0.0225	0.9446	0.1361
	부동산소득액	1,342.4	1,216	2,155.3	2,634.7	-812.9	-1.16	0.2602	0.796	0.0479
	이전소득액	2,279.2	5,935.3	1,737.8	4,124.7	541.4	0.51	0.609	0.3233	0.4553
	기타소득액	1,376.7	4,599.3	1,343.3	3,637.6	33.5	0.04	0.9693	0.3278	0.6567
2차	근로소득액	5,292.8	3,256.1	4,870.8	2,830.5	422	1.09	0.2769	0.4459	0.2698
	금융소득액	628.5	1,363.5	648.5	1,469.3	-190.62	-0.08	0.9397	0.8264	0.8812
	부동산소득액	3,306.7	4,441.2	1,844.3	2,745.6	1,462.4	1.06	0.2998	0.394	0.3301
	이전소득액	333.2	161.4	205.6	195.9	127.6	1.59	0.1293	0.1284	0.5667
	기타소득액	1,354.9	3,119.6	1,281.8	5,745.9	73.1	0.07	0.9417	0.2066	0.6442

주 : 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

2) Pr>|Z|는 비모수검정(nonparametric test)을 통해 계산된 값임.

이전소득액은 1차년도와 2차년도 간 평균 차이가 가장 큰 편으로 나타났다. <표 3-26>의 이전소득액에 대한 응답가구수 분포를 보면, 이전소득액이 있는 가구수가 1차년도의 경우 48가구였으나 2차년도에는 약 77% 대 폭 감소한 11가구로 나타났다. 추가로 이전소득액의 세부 항목에 대한 응답가구수를 살펴보면, 따로 사는 부모님 도움 및 따로 사는 자녀 도움의 경우 CAPI와 CAWI 모두 가장 큰 차이를 보였다. 이러한 결과는 실제 소득이 발생하지 않았거나, 설문조사의 학습효과로 인한 영향으로 볼 수 있지만 직접적으로 판단하기는 쉽지 않다. 하지만 조사도구와는 상관없다고 볼 수 있다.

추가적으로 금융자산, 지난 한 달 소득과 월평균 저축액의 분포에 대해서도 살펴보았다. 먼저 이상치를 제거하지 않고 검증한 결과인 <표 3-27>를 보면, 모든 변수들이 1차년도와 2차년도에서 조사도구 간 통계적으로 유의한 차이가 없다고 나타났다. 분산 검증에서도 CAPI와 CAWI 간의 분산은 통계적으로 유의한 차이가 없다고 나타났다.

〈표 3-26〉 이전소득액 및 이전소득액의 세부 항목별 응답가구수

(단위: 가구)

		CAPI		CAWI	
		1차	2차	1차	2차
이전소득액		48	11	48	11
세부항목	국민기초생활보호대상자 지원금	0	0	6	3
	기타 정부보조금	8	8	9	5
	사회단체 보조금	0	0	1	1
	따로 사는 부모님 도움	36	1	35	0
	따로 사는 자녀의 도움	12	0	3	0
	그 외 친척, 친지 보조금	1	0	3	3
	기타	0	2	0	2

주: 이상치를 제거하지 않고 분석함.

〈표 3-27〉 금융자산 및 기타변수 차이 검증

		CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t	Pr> Z  <sup>2)</sup>	Levene 검정
		평균 (만 원)	S.D	평균 (만 원)	S.D					
1차	금융자산	7,993.4	18,666.8	7,295.8	13,598.5	697.6	0.32	0.7512	0.2378	0.4589
	지난 한 달 소득	469.6	440.4	395.5	278.5	74.1133	1.48	0.1412	0.167	0.2655
	월평균 저축액	108.4	89.7524	100.0	132.1	8.436	0.6	0.5465	0.1775	0.4918
2차	금융자산	8,082.3	13,366.6	5,986.6	12,931.1	2,095.7	1.29	0.1963	0.0003	0.922
	지난 한 달 소득	467.9	343.6	476.8	676.5	-8.9022	-0.14	0.8848	0.3812	0.4059
	월평균 저축액	110.3	102.2	114	153.7	-3.7142	-0.23	0.819	0.3803	0.3092

주: 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

2) Pr&gt;|Z|는 비모수검정(nonparametric test)을 통해 계산된 값임.

다음으로 이상치의 개수는 모든 항목에서 CAPI보다 CAWI에서 더 많이 발생하였다(표 3-28 참조). 2차년도 금융자산의 경우 CAPI에서 9가구, CAWI에서 53가구로 나타났는데, 이에 대한 이상치 빈도를 보면 7만 원이 전체의 대부분을 차지하고 있다(CAPI: 8가구, CAWI: 51가구). 이 가구들에 대한 금융자산의 세부항목을 살펴보면 첫 세부항목인 은행예금만 있고, 나머지 항목(주식/채권/신탁, 저축성 보험, 아직 타지 않은 계, 개인

적으로 다른 사람에게 빌려 준 돈)은 모두 없다고 응답하였다. 실제값(true value)일 수 있지만, 보통 설문조사 시 5만 원보다는 많고 10만 원보다 적은 7만 원으로 응답하는 경향도 반영되었다고 볼 수 있어 판단하기 쉽지 않다.

다음 <표 3-29>는 이상치 제거 후의 검증 결과로, 1차년도 월평균 저축액 평균이 조사도구 간 통계적으로 유의한 차이를 가지고 있다. 이를 제외한 나머지 검증 결과는 이상치 제거 전과 동일하게 나타났다.

<표 3-28> 금융자산 및 기타변수에 대한 이상치 개수

(단위: 가구)

		CAPI	CAWI
1차	금융자산	1	3
	지난 한 달 소득	1	4
	월평균 저축액	1	4
2차	금융자산	9	53
	지난 한 달 소득	0	4
	월평균 저축액	0	4

<표 3-29> 금융자산 및 기타변수 차이 검증: 이상치 제거 후 분석

		CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t	Pr> Z  <sup>2)</sup>	Levene 검정
		평균 (만 원)	S.D	평균 (만 원)	S.D					
1차	금융자산	6,257.1	8,477.7	6,788.1	11,201.9	-531.0	-0.43	0.6653	0.3271	0.3105
	지난 한 달 소득	432.7	261.5	387.1	197.7	45.6171	1.47	0.1432	0.2639	0.1252
	월평균 저축액	109.7	89.5	88.2	77.0627	21.4859	1.98	0.0486	0.0898	0.2392
2차	금융자산	7,880.9	10,604.5	6,773.3	10,139.3	1,107.6	0.8	0.425	0.0654	0.8582
	지난 한 달 소득	467.9	343.6	418.8	286.7	49.058	1.18	0.2413	0.3724	0.4402
	월평균 저축액	110.3	102.2	103.4	110.5	6.8325	0.48	0.6344	0.3707	0.7862

주: 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미함.

2) Pr>|Z|는 비모수검정(nonparametric test)을 통해 계산된 값임.

#### 다. 만족도 및 건강상태

개인응답자 설문 문항에는 주관적 응답을 받는 만족도 변수와 건강상태 관련 변수들이 존재한다. 이에 해당되는 조직몰입도, 직무만족도, 요인별 직무만족도, 생활만족도, 건강 관련 변수의 CAPI와 CAWI 집단 간 응답 차이를 분석하였다. 이들 만족도 문항은 리커트 5점 척도에 의해 조사되며, 분석방법으로 두 집단 평균차 검정 분석을 실시하였다.

<표 3-30>에서 보듯이 1차년도에 이어 2차년도 조사에서 전체 21개 항목으로 구성된 직장만족도 대부분 변수들은 CAPI와 CAWI 두 집단 간 평균점수 차이가 5% 수준에서 통계적으로 유의하게 나타났다. 조직몰입도 5개 변수를 살펴보면 ‘지금 직장은 다닐 만하다’와 ‘이 직장을 계속 다니고 싶다’ 항목의 1차조사에서는 CAPI와 CAWI 두 집단 간 평균점수가 통계적으로 유의하지 않은 것으로 나타났으나, 2차조사에서는 CAWI 집단 평균점수는 CAPI 집단의 평균점수보다 통계적으로 유의하게 낮은 것으로 분석되었다. 2개 문항을 제외한 3개 문항은 1차 조사와 마찬가지로 2차 조사에서도 5% 수준에서 유의한 것으로 나타났다. 직무만족도 5개 변수 중 ‘이 일을 계속하고 싶다’ 항목만 1차와 2차조사에서 다른 양상을 보인 반면 4개 변수는 1차조사와 2차조사에서 CAWI 집단 평균점수는 CAPI 집단의 평균점수보다 통계적으로 유의하게 낮은 것으로 나타났다. 7가지 요인별 직무만족도를 살펴보면 1차 조사에서 임금 및 보수 항목을 제외한 나머지 항목은 통계적으로 유의한 차이가 있는 것으로 나타났으며, 2차 조사에서는 근로시간 항목을 제외한 나머지 항목에서 CAPI 집단보다 CAWI 집단의 평균점수가 높게 나타난 것으로 보인다. 전반적 일자리 만족도와 전반적인 일 만족도는 1차년도와 2차년도에서 CAPI 집단보다 CAWI 집단의 평균점수가 높게 나타났다. 이는 주관적으로 응답받는 직장만족도 변수들은 CAPI와 CAWI 조사방식의 변화로 인한 모드효과<sup>20)</sup>가 존재한다는 사실을 다시 한번 확인하였다.

20) 모드효과란 조사방식(면접조사, 전화조사, 우편조사, 인터넷조사 등)의 변화로 인해서 발생하는 표본이탈, 무응답오차, 측정오차 등을 종합적으로 일컫는 용어이다(최효미 외, 2012; 김유빈 외, 2017).

〈표 3-30〉 직장만족도 응답 차이 검증

	1차			2차		
	CAPI 평균	CAWI 평균	Pr> t	CAPI 평균	CAWI 평균	Pr> t
조직몰입도						
지금 직장은 다닐 만하다	3.6707	3.4716	0.1013	3.6543	3.4269	0.0345
이 직장에 들어온 것을 기쁘게 생각한다	3.6463	3.3920	0.0353	3.6420	3.3509	0.0209
친구에게 이 직장을 추천하고 싶다	3.4390	3.1023	0.0149	3.5185	3.1228	0.0047
이 직장을 다른 사람들에게 자랑할 수 있다	3.5122	3.1932	0.0164	3.4815	3.2281	0.0429
이 직장을 계속 다니고 싶다	3.7439	3.5455	0.1212	3.8025	3.5380	0.0234
직무만족도						
현재 하고 있는 일에 만족하고 있다	3.6765	3.3058	0.0002	3.6061	3.3350	0.0049
이 일을 열정적으로 하고 있다	3.6863	3.3301	0.0010	3.5859	3.3695	0.0301
이 일을 즐겁게 하고 있다	3.7059	3.2330	<.0001	3.5859	3.2857	0.0029
이 일을 보람을 느끼면서 하고 있다	3.7451	3.2621	<.0001	3.5960	3.2611	0.0010
이 일을 계속하고 싶다	3.8529	3.4272	<.0001	3.6667	3.4877	0.1184
요인별 직무만족도						
임금/보수	2.9314	3.0777	0.1816	2.7980	2.9901	0.0819
취업안정성	2.5000	2.7816	0.0126	2.4747	2.7241	0.0391
일의 내용	2.3922	2.7039	0.0043	2.3535	2.6108	0.0080
근무환경	2.4216	2.7427	0.0031	2.4848	2.7635	0.0102
근로시간	2.3922	2.7282	0.0038	2.5152	2.6700	0.1935
개인발전가능성	2.5588	2.8932	0.0025	2.5657	2.9606	0.0002
의사소통 및 인간관계	2.3627	2.7573	0.0003	2.4747	2.7438	0.0146
인사고과의 공정성	2.7805	3.0057	0.0698	2.6914	3.0643	0.0015
복지후생제도	2.9512	3.2159	0.0557	2.8519	3.1930	0.0170
전반적 일자리만족도	2.4902	2.8058	0.0024	2.4545	2.7931	0.0004
전반적 일만족도	2.3922	2.7864	0.0001	2.3939	2.7192	0.0003

주: 1) 조직몰입도, 직무만족도는 (1)매우 불만족~(5)매우 만족, 요인별 직무만족도와 전반적 만족도는 (1)매우 만족~(5)매우 불만족이므로 주의.

2) 1차, 2차년도 균형패널 사용.



<표 3-31>에서는 생활만족도 차이를 검증한 결과, 생활만족도 7개 항목 모두 1차년도에 이어 2차년도 조사에서도 CAPI 집단의 평균점수는 CAWI 집단의 평균점수보다 통계적으로 유의하게 낮은 것으로 분석되었다.

<표 3-32>는 1차년도와 2차년도의 건강상태 변수의 차이를 검증하고 있다. 우선 1차 조사에서는 ‘1년 대비 건강상태’ 변수를 제외한 다른 항목에서 CAPI 집단보다 CAWI 집단의 평균점수가 높은 것으로 나타나 통계적으로 유의한 결과를 보였다. 부연하면 현재 건강상태는 CAWI 집단이 CAPI 집단보다 조금 더 건강하다는 응답에 체크한 것으로 보인다. 또한 감각기관 장애, 육체적 제약, 활동제약 1(배우기, 기억하기, 집중하기), 활동제약 3(쇼핑하기, 병원가기 등), 활동제약 4(작업활동 어려움)는 CAWI 집단보다 CAPI 집단의 평균점수가 높게 나타나 통계적으로 유의한 것으로 분석되었다. 2차 조사에서는 건강상태 및 장애 응답 항목에 대하여 CAWI 집단 평균점수는 CAPI 집단의 평균점수보다 통계적으로 유의하게 높은 것으로 분석되었다.

<표 3-31> 생활만족도 차이 검증

		CAPI		CAWI		diff1)	t-value	Pr> t
		평균	S.D.	평균	S.D.			
1차	가족수입	2.9071	0.8126	3.1111	0.8126	-0.204	-2.42	0.0158
	여가생활	2.8286	0.8477	3.0968	0.8782	-0.2682	-2.98	0.0030
	주거환경	2.5929	0.7384	2.9391	0.8353	-0.3462	-4.16	<.0001
	가족관계	2.2786	0.8053	2.4803	0.7812	-0.2017	-2.47	0.0140
	친인척관계	2.5857	0.8395	2.8602	0.7763	-0.2745	-3.32	0.0010
	사회적 친분	2.4143	0.6993	2.7025	0.7053	-0.2882	-3.96	<.0001
	전반적 만족도	2.5357	0.6394	2.8029	0.7495	-0.2672	-3.8	0.0002
2차	가족수입	2.5214	0.6621	2.8172	0.7288	-0.2958	-4.04	<.0001
	여가생활	2.8714	0.7942	3.1613	0.8391	-0.2899	-3.39	0.0008
	주거환경	2.8071	0.8724	3.1147	0.8821	-0.3076	-3.38	0.0008
	가족관계	2.6429	0.8142	2.9140	0.8178	-0.2711	-3.21	0.0015
	친인척관계	2.2571	0.7333	2.4910	0.8434	-0.2339	-2.79	0.0054
	사회적 친분	2.5071	0.7911	2.8280	0.8174	-0.3208	-3.83	0.0001
	전반적 만족도	2.4214	0.7401	2.7419	0.6928	-0.3205	-4.37	<.0001

주 : 1) 생활만족도는 (1)매우 만족~(5)매우 불만족인 5점 척도.

2) 1차, 2차년도 균형패널 사용.

〈표 3-32〉 건강상태 및 장애 응답 차이 검증

	1차			2차		
	CAPI 평균	CAWI 평균	Pr> t	CAPI 평균	CAWI 평균	Pr> t
현재 건강상태	2.4071	2.6918	0.0007	2.4286	2.7706	<.0001
1년 전 대비 건강상태	3.0071	3.0645	0.4046	2.9857	3.0824	0.0923
보통사람 대비 건강상태	2.6571	2.8495	0.0094	2.6571	2.8746	0.0023
감각기관장애	2.0000	1.9391	<.0001	1.9929	1.9677	0.05
육체적 제약	1.9786	1.8961	0.0002	1.9571	1.9032	0.0296
활동제약1 (배우기, 기억하기 등)	1.9857	1.9068	0.0001	1.9786	1.9355	0.0252
활동제약2 (옷입기, 목욕하기 등)	2.0000	1.9749	0.0079	1.9857	1.9821	0.7848
활동제약3 (쇼핑하기, 병원가기 등)	2.0000	1.9677	0.0026	1.9857	1.9534	0.0462
활동제약4 (작업활동 어려움)	1.9571	1.9068	0.0404	1.9571	1.9176	0.0974

주: 1) 건강상태는 (1)매우 건강~(5)매우 불건강인 5점 척도. 장애 및 활동제약 항목은 (1)예 (2)아니오임.

2) 1차, 2차년도 균형패널 사용

〈표 3-33〉에서는 건강 관련 변수로 흡연과 음주 변수들을 비교해 보았다. 흡연 여부는 ‘흡연/비흡연/과거에 흡연한 적 있다’로 응답하게 되어 있는 변수이며, 하루 평균 흡연량은 개수를 연속형으로 질문한 것이 아니라 구간을 나누어 질문하였다. 금연 시도는 ‘시도한 적 있다/없다’로 구분해 답하도록 되어 있다. 음주 관련 질문들도 흡연과 유사하게 설계되어 있다(김유빈 외, 2017). 이와 같이 개인적으로 민감한 흡연 및 음주와 관련된 문항의 경우 CAPI와 CAWI의 분포를 비교해 보고 그 분포에 차이가 있는지를 살펴본다. 흡연 여부, 음주 여부 문항에 대해서는 카이제곱 검정을 실시하였으며, 금연과 금주 시도 문항에 대해서는 Fisher의 정확 검정을 적용하였다. 1차년도 조사에서는 CAPI와 CAWI 조사도구 방식에 있어 흡연 여부 문항만 차이가 있고, 다른 변수는 차이가 없는 것으로 나타났다. 2차년도 조사에서는 흡연 여부 및 금주 시도 문항만 통계적으로 유의하게 차이가 있는 것으로 나타났다.

〈표 3-33〉 흡연 및 음주 변수의 차이 검증

			CAP	CA.WI	Pr <= P
1차	흡연여부	피운다	14(10.0)	57(20.4)	0.0019
		과거에 피웠으나 현재 피우지 않는다	26(18.6)	71(25.5)	
		피워본 적 없다	100(71.4)	151(54.1)	
	금연시도	있음	4(28.6)	26(45.6)	0.3668
		없음	10(71.4)	31(54.4)	
	음주여부	마신다	93(66.4)	184(66.0)	0.9941
		과거에 마셨으나 현재 마시지 않는다	19(13.6)	38(13.6)	
		마셔본 적 없다	28(20.0)	57(20.4)	
	금주시도	있다	3(3.2)	9(4.9)	0.7563
		없다	90(96.8)	175(95.1)	
2차	흡연여부	피운다	13(9.3)	62(22.2)	0.0036
		과거에 피웠으나 현재 피우지 않는다	32(22.9)	63(22.6)	
		피워본 적 없다	95(67.9)	154(55.2)	
	금연시도	있다	5(38.5)	27(43.6)	1.0000
		없다	8(61.5)	35(54.4)	
	음주여부	마신다	90(64.3)	187(67.0)	0.4598
		과거에 마셨으나 현재 마시지 않는다	17(12.1)	40(14.3)	
		마셔본 적 없다	33(23.6)	52(18.6)	
	금주시도	있다	1(1.1)	17(9.1)	0.0090
		없다	89(98.9)	170(90.9)	

주: 1) Pr <= P는 카이제곱검정 및 Fisher의 정확검정을 통해 계산된 값임.

2) 1차, 2차년도 균형패널 사용.

## 라. 일자리 변수

여기에서는 종사상 지위, 통계기준에 따른 근로형태별 비정규직 근로자, 평균근로시간 등 일자리 관련 변수의 조사방법에 따른 평균 및 분포 차이를 비교한다. <표 3-34>는 조사방법에 따른 종사상 지위별 분포 차

이를 비교한 것인데 비임금근로자의 경우 종사상 지위별로 표본수가 충분치 않아 분포 차이의 검정은 Fisher의 정확한 검정(Fisher's exact test)으로 나타났다. 검정 결과, 차수별 임금근로자와 비임금근로자 모두 종사상 지위 분포의 조사방법에 따른 차이가 유의하지 않은 것으로 나타났다.

<표 3-35>는 임금근로자 중 비정규직 근로자 비중의 조사방법에 따른 차이를 비교한 것이다. 비정규직 비중은 통계 기준과 자기선언 기준에 따라 각각 계산되었다. 여기서 통계 기준은 2002년 노사정위원회에서 합의한 바에 따른 것이며, 자기선언 기준은 응답자의 주관적 판단에 따른 것이다. 자기선언적 비정규직 근로자 비중은 응답을 거부한 표본을 제외하고 계산하였다.

CAWI의 통계 기준 비정규직 비율은 1차 63.6%, 2차 65.5%로 CAPI의

<표 3-34> 종사상 지위 분포의 조사방법에 따른 차이 비교

		종사상 지위	CAPI		CAWI		Pr<=P
			표본수	백분율	표본수	백분율	
1차	임금 근로자	상용직	71	86.7	162	92.1	0.367
		임시직	8	9.8	11	6.3	
		일용직	3	3.7	3	1.7	
		전 체	82	100.0	176	100.0	
	비임금 근로자	고용주	6	30.0	12	36.0	0.730
		자영업자	14	70.0	17	62.0	
		무급가족 종사자	0	0.0	1	2.0	
		전 체	20	100.0	30	100.0	
2차	임금 근로자	상용직	72	88.9	154	90.1	0.367
		임시직	7	8.6	12	7.0	
		일용직	2	2.5	5	2.9	
		전 체	81	100.0	171	100.0	
	비임금 근로자	고용주	5	27.8	15	46.9	0.237
		자영업자	13	72.2	17	53.1	
		무급가족 종사자	0	0.0	0	0.0	
		전 체	18	100.0	32	100.0	

2배에 달하는 것으로 나타났다. 이러한 차이는 한시적 근로자와 비전형근로자 중 특수형태근로종사자의 비중 차이에 기인한다. 한시적 근로자 비중은 1차 CAPI에서 31.7%, CAWI에서 54.6%, 2차 CAPI에서 27.2%, CAWI에서 55.6%인 것으로 나타났고, 특수형태근로종사자는 각각 1차 4.9%와 13.1%, 2차 4.9%와 11.7%로 나타났다.

한시적 근로자 중에서도 특히 기간제 근로자(계약기간을 정한 근로자) 비중은 1차 조사에서 CAPI와 CAWI 각각 17.1%, 40.9%로 나타나 그 차이가 22.8%p로 높게 나타났다. 근로기간의 정함 여부로 한시적 근로자는 기간제와 반복갱신 및 기대불가로 나뉘는데 응답자들의 해당 문항에 대한 오해로 인하여 기간제에서만 조사방법에 따른 응답률 차이가 나타난 것으로 판단하였다. 따라서 2차 조사에서는 설문문항에 ‘근로기간의 정함은 근로계약서상에 특정 일자나 기간이 명시된 것을 의미하는 것으로, 정년은 이에 해당되지 않습니다.’라고 부가적인 설명을 표기하였으나 2차 조사에서도 여전히 그 비중은 CAPI가 16.1%, CAWI는 44.4%로 차이가 크게 나타났다(표 3-36 참조).

한편 자기선언 기준 비정규직 비율은 반대로 1차 CAPI 21.0%, CAWI 12.6%, 2차 CAPI 18.8%, CAWI 14.1%로 CAPI가 CAWI에 비해 높았으나 1차에서만 그 차이가 유의한 것으로 나타났다.

〈표 3-35〉 임금근로자 중 비정규직 근로자 비중의 조사방법에 따른 차이 비교

		CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t
		비중	S.D.	비중	S.D.			
1차	통계기준	0.3659	0.4846	0.6364	0.4824	-0.2705	-4.19	0.0000
	자기선언기준	0.2099	0.4098	0.1257	0.3325	0.0842	1.75	0.0820
	불일치여부 <sup>2)</sup>	0.1975	0.4006	0.5314	0.5004	-0.3339	-5.27	0.0000
2차	통계기준	0.3457	0.4786	0.6550	0.4768	-0.3093	-4.80	0.0000
	자기선언기준	0.1875	0.3928	0.1412	0.3492	0.0463	0.94	0.3484
	불일치 여부 <sup>2)</sup>	0.1750	0.3824	0.5118	0.5013	-0.3368	-5.32	0.0000

주: 1) diff는 각 차수별 CAPI와 CAWI의 비정규직 근로자 비중 차이를 의미.

2) 통계 기준과 자기선언 기준에 따른 비정규직 여부가 일치하지 않는 경우.

〈표 3-36〉 비정규직 근로형태별 조사방법에 따른 차이 비교(통계기준)

		CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t
		평균	S.D.	평균	S.D.			
1차	비정규직	0.3659	0.4846	0.6364	0.4824	-0.2705	-4.19	0.00
	한시적 근로	0.3171	0.4682	0.5455	0.4994	-0.2284	-3.49	0.00
	기간제	0.1707	0.3786	0.4091	0.4931	-0.2384	-3.88	0.00
	반복갱신	0.0854	0.2811	0.1136	0.3183	-0.0283	-0.69	0.49
	기대불가	0.0610	0.2408	0.0227	0.1495	0.0382	1.56	0.12
	비전형 근로	0.1098	0.3145	0.1761	0.3820	-0.0664	-1.37	0.17
	파견	0.0122	0.1104	0.0057	0.0754	0.0065	0.55	0.58
	용역	0.0366	0.1889	0.0114	0.1063	0.0252	1.37	0.17
	특수형태근로	0.0488	0.2167	0.1307	0.3380	-0.0819	-2.01	0.05
	가정 내 근로	0.0000	0.0000	0.0170	0.1298	-0.0170	-1.19	0.24
	일일근로	0.0366	0.1889	0.0284	0.1666	0.0082	0.35	0.73
	시간제 근로	0.0854	0.2811	0.0682	0.2528	0.0172	0.49	0.62
	비정규직	0.3457	0.4786	0.6550	0.4768	-0.3093	-4.80	0.00
	한시적 근로	0.2716	0.4476	0.5556	0.4984	-0.2840	-4.36	0.00
2차	기간제	0.1605	0.3694	0.4444	0.4984	-0.2840	-4.57	0.00
	반복갱신	0.0617	0.2422	0.0936	0.2921	-0.0318	-0.85	0.40
	기대불가	0.0494	0.2180	0.0175	0.1317	0.0318	1.44	0.15
	비전형 근로	0.1111	0.3162	0.1871	0.3912	-0.0760	-1.53	0.13
	파견	0.0123	0.1111	0.0175	0.1317	-0.0052	-0.31	0.76
	용역	0.0247	0.1561	0.0058	0.0765	0.0188	1.29	0.20
	특수형태근로	0.0494	0.2180	0.1170	0.3223	-0.0676	-1.71	0.09
	가정 내 근로	0.0000	0.0000	0.0351	0.1845	-0.0351	-1.71	0.09
	일일근로	0.0247	0.0173	0.0409	0.1987	-0.0162	-0.65	0.52
	시간제 근로	0.0864	0.2827	0.0760	0.2658	0.0104	0.28	0.78

주: 1) diff는 각 차수별 CAPI와 CAWI의 비정규직 근로자 비중 차이를 의미.

〈표 3-37〉은 주당 평균근로시간을 조사방법에 따라 비교한 것이다. 주당 평균노동시간에 대한 무응답은 1, 2차와 두 조사방법 모두에서 발견되지 않았다. 임금근로자의 주당 평균근로시간은 정규근로시간이 정해져 있는 경우 정규근로시간과 초과근로시간의 합으로, 정해져 있지 않은 경우 주당 평균근로시간으로 측정하였다. 1차 임금근로자의 경우에만 조사방법에 따른 주당 평균근로시간 차이만 유의한 것으로 나타났다.

〈표 3-37〉 평균근로시간의 조사방법에 따른 차이 비교

		CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t
		평균(시간)	S.D.	평균(시간)	S.D.			
1차	임금근로자	42.6	9.93	45.9	13.04	-3.3	-2.04	0.0427
	비임금근로자	42.3	17.07	44.2	16.41	-2.0	0.41	0.6822
2차	임금근로자	43.3	10.42	43.0	11.15	0.4	0.25	0.7997
	비임금근로자	43.7	17.90	43.8	16.35	-0.1	0.03	0.9768

주: 1) diff는 각 차수별 CAPI와 CAWI의 평균근로시간 차이를 의미.

#### 마. 개인평균소득

〈표 3-38〉은 차수별 임금근로자와 비임금근로자의 평균소득 유무의 조사방법에 따른 차이를 보여준다. 특히 비임금근로자 중 적자를 보고 있는 경우와 무급가족종사자는 평균소득이 없는 것으로 간주하여 분석하였다. CAPI에서는 임금근로자 중 평균소득이 없다고 응답한 경우가 없었음에도 불구하고, 임금근로자와 비임금근로자 모두 평균소득 없음의 비중 차이가 유의하지 않은 것으로 나타났다. 반면 2차 조사에서는 CAPI와 CAWI의 비임금근로자 평균소득 유무 차이가 유의한 것으로 나타나고 있다.

〈표 3-39〉은 임금근로자와 비임금근로자 평균소득의 조사방법에 따른 차이를 비교한 것이다. 임금근로자의 평균소득은 월평균임금으로, 비임금근로자는 월평균소득으로 계산되었으며, 비임금근로자 중 적자를 보

〈표 3-38〉 평균소득 유무의 조사방법에 따른 차이 비교

		CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t
		평균소득 없음	S.D.	평균소득 없음	S.D.			
1차	임금근로자	-	-	0.0114	0.1063	-0.0114	-0.97	0.3344
	비임금근로자	0.0500	0.2236	0.2069	0.4123	-0.1569	-1.55	0.1281
2차	임금근로자	-	-	-	-	-	-	-
	비임금근로자	-	-	0.1563	0.3689	-0.1563	-1.79	0.0799

주: 1) diff는 각 차수별 CAPI와 CAWI의 평균소득 없음의 비중 차이를 의미.

〈표 3-39〉 평균소득의 조사방법에 따른 차이 비교

		CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t
		평균소득(만 원)	S.D.	평균소득(만 원)	S.D.			
1차	임금근로자	284.2	132.1	292.8	249.7	-8.6	-0.29	0.7702
	비임금근로자	279.5	401.7	273.8	401.7	5.7	0.06	0.9549
2차	임금근로자	324.7	159.3	313.9	430.3	10.8	0.22	0.8269
	비임금근로자	247.8	215.6	296.8	324.9	-49.0	-0.57	0.5715

주: 1) diff는 CAPI와 CAWI의 평균소득 차이를 의미.

고 있는 경우와 무급가족종사자의 평균소득은 '0'으로 간주하였다. 모수 검정에서는 1, 2차 모두 평균소득의 조사방법에 따른 차이가 유의하지 않은 것으로 나타났으나, 비모수 검정에서는 2차 임금근로자의 평균소득 차이가 유의한 것으로 나타났다.

#### 4. 항목무응답률

##### 가. 교육, 혼인상태 변수

개인설문에서 응답자가 응답하기 꺼리는 민감한 질문에 대한 무응답률을 살펴보고자 한다. 개인설문에서 모름/무응답 버튼을 허용한 작년 한 해 근로소득변수를 제외하고 거의 대부분 변수는 무응답이 없는 것으로 확인되었다. 다만 혼인상태 관련 변수에서 1차년도와 2차년도 조사에서 동일한 시스템적 로직 오류가 있는 것으로 확인되었다.

<표 3-40>에서 개인용 자료의 주요 변수 중 재학 중인 개인의 교육과 혼인상태에 관한 무응답률을 살펴본다. 신규조사자는 과거에 다녔거나 다니고 있는 고등학교/대학교/대학원에 대한 교육력 설문이 진행된다. 한국노동패널조사에서 신규로 들어오는 응답자는 개인용 설문에 응답해야 하는 문항이 상당하다. 더욱이 교육력과 관련된 문항은 꺼려할 수도 있다. 그러나 표에서 제시된 바와 같이 1차년도와 2차년도 조사에서 응답자의 현재 재학 중인 학교의 유형에 대한 무응답은 나타나지 않았다.



〈표 3-40〉 재학 여부, 혼인상태 무응답률(CAPI vs CAWI)

	1차				2차			
	CAPI		CAWI		CAPI		CAWI	
	무응답률		무응답률		무응답률		무응답률	
	리뷰 전	리뷰 후	리뷰 전	리뷰 후	리뷰 전	리뷰 후	리뷰 전	리뷰 후
재학 중인 학교의 유형	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
혼인상태	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
혼인상태 변화 여부	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
혼인상태 변화 횟수	91.6	0.0	89.1	0.0	93.2	0.0	87.7	0.0
혼인상태 변화내용 1	91.6	0.0	89.1	0.0	93.2	0.0	87.7	0.0
혼인상태 변화내용 2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
혼인상태 변화내용 3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
혼인상태 변화내용 4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

주: 1) 리뷰 후 자료는 기존 CAPI 버전의 프로그램을 CAWI(웹) 버전으로 변경하는  
과정 중 발생한 로직 에러에 따른 오류로 판단되어 자료 수정된 수치임.

2) 1차, 2차년도 균형패널 사용.

한편 혼인력과 관련된 변수의 무응답률은 로직 오류로 인하여 확인할 수 있었다. 혼인상태가 미혼(한 번도 결혼한 적 없다)의 경우를 제외하고는 혼인상태 변화가 있었거나 없었던 개인에 대해 변화 내용-1에서 처음 결혼함으로 응답되어야 하나, 로직 오류로 인하여 혼인상태 변화가 없는 개인에게 응답받지 못한 케이스를 발견한 것이다. 리뷰 후에는 해당 문항에 대한 응답값이 전부 채워진 결과, 혼인상태 관련 변수의 무응답률이 발견되지 않았다.

#### 나. 개인 근로소득변수

1차와 2차년도 웹조사 실험설계 시 개인설문에서 작년 한 해 근로소득 변수에 한정하여 모름/무응답 버튼을 허용하였으며, 해당 차수에 조사된 개인 근로소득변수에 대한 무응답률을 확인하고자 한다. <표 3-41>은 개인의 작년 한 해 근로소득 관련 문항의 t-test 결과이다. 세전/세후 근로

소득 문항 t-test에 사용된 자료는 소득액이 있다는 응답자의 소득액 문항의 무응답 비중이다. 즉 무응답률의 결과가 문항에 대한 응답 빈도에 차이가 있는가를 검증한 것이다. t-test 결과, 1차년도와 2차년도의 세전 근로소득액과 세후 근로소득액은 조사도구에 따라 유의한 차이를 갖는 것으로 나타났다.

〈표 3-41〉 작년 한 해 근로소득 무응답률의 집단 간 차이

		CAPI		CAWI		diff1)	t-value	Pr> t
		평균	S.D.	평균	S.D.			
1차	근로소득 유무	0.0000	0.0000	0.0000	0.0000	-	-	-
	세전 근로소득	0.0286	0.2382	0.1864	0.5824	-0.1578	-3.92	0.0001
	세후 근로소득	0.0857	0.5016	0.3763	0.9955	-0.2906	-3.97	<.0001
2차	근로소득 유무	0.0000	0.0000	0.0000	0.0000	-	-	-
	세전 근로소득	0.0143	0.1690	0.1505	0.5286	-0.1363	-3.92	0.0001
	세후 근로소득	0.0214	0.2535	0.3548	0.9706	-0.3334	-5.38	<.0001

주: 1) diff는 CAPI 평균에서 CAWI 평균을 뺀 차이를 의미.

2) 1차, 2차년도 균형패널 사용.

#### 다. 일자리 관련 변수

여기서는 직업력 자료 중 응답자가 모름/무응답을 선택할 수 있는 변수들의 무응답률을 조사방법에 따라 비교한다. 직업력 자료 중 일자리 시작 시기, 자기선언적 비정규직 근로자 여부, 종업원수, 평균소득은 응답자가 모름/무응답을 선택할 수 있다.

〈표 3-42〉는 일자리 시작 시기의 무응답률을 조사방법에 따라 비교한 것이다. 검정 결과, 조사방법에 따른 차수별 일자리 시작 시기 무응답률 차이는 유의하지 않았다.

〈표 3-43〉는 임금근로자 중 자기선언적 비정규직 근로자 여부와 종업원수 무응답률을 조사방법에 따라 비교한 것이다. 비임금근로자의 종업원수 무응답률은 표본수가 적어 분석에서 제외하였다. 검정 결과, 차수별 조사방법에 따른 자기선언적 비정규직 근로자 무응답률과 종업원수 무응

답률 차이는 유의하지 않았다.

<표 3-44>는 차수별 평균소득 무응답률의 조사방법에 따른 차이를 비교한 것이다. 임금근로자의 경우 1차에서만 평균소득 무응답률의 유의한 차이가 발견되었고 2차에는 임금근로자와 비임금근로자 모두 무응답이 나타나지 않았다. 앞서 평균소득의 조사방법에 따른 차이가 유의하지 않

<표 3-42> 일자리 시작 시기(연/월/일) 무응답률의 집단 간 비교

		일자리 시작 시기	CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t
			무응답	S.D.	무응답	S.D.			
1차	임금 근로자	연	-	-	-	-	-	-	-
		월	-	-	-	-	-	-	-
		일	-	-	-	-	-	-	-
	비임금 근로자	연	-	-	0.0333	0.1826	0.033	-0.81	0.4198
		월	-	-	0.0333	0.1826	0.033	-0.81	0.4198
		일	-	-	0.0333	0.1826	0.033	-0.81	0.4198
2차	임금 근로자	연	-	-	-	-	-	-	-
		월	-	-	-	-	-	-	-
		일	-	-	-	-	-	-	-
	비임금 근로자	연	0.0556	0.2357	0.0938	0.2961	-0.0382	-0.47	0.6410
		월	0.0556	0.2357	0.0938	0.2961	-0.0382	-0.47	0.6410
		일	0.0556	0.2357	0.0938	0.2961	-0.0382	-0.47	0.6410

주: 1) diff는 CAPI와 CAWI의 무응답률 차이를 의미.

<표 3-43> 임금근로자 중 비정규직 여부와 종업원수 무응답률의 조사방법에 따른 차이 비교

		CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t
		무응답	S.D.	무응답	S.D.			
1차	비정규직 여부	0.0122	0.1104	0.0057	0.0754	0.0065	0.55	0.5803
	종업원수	0.0441	0.2069	0.1019	0.3035	-0.0578	-1.43	0.1536
2차	비정규직 여부	0.0123	0.1111	0.0058	0.0765	0.0065	0.54	0.5890
	종업원수	0.0571	0.2338	0.1218	0.3281	-0.0647	-1.49	0.1384

주: 1) diff는 CAPI와 CAWI의 무응답률 차이를 의미.

〈표 3-44〉 평균소득 무응답률의 조사방법에 따른 차이 비교

		CAPI		CAWI		diff <sup>1)</sup>	t-value	Pr> t
		무응답	S.D.	무응답	S.D.			
1차	임금근로자	-	-	0.0511	0.0511	-0.0511	-2.09	0.0372
	비임금근로자	-	-	0.0333	0.1826	-0.0333	-0.81	0.4198
2차	임금근로자	-	-	-	-	-	-	-
	비임금근로자	-	-	-	-	-	-	-

주: 1) diff는 CAPI와 CAWI의 무응답률 차이를 의미.

았음을 볼 때, 이는 CAPI에서 조사자가 대략적인 평균소득을 조사해 오는 반면에 CAWI에서는 응답자가 부정확한 답변을 피하려는 성향을 보이기 때문으로 생각된다. 보다 정확한 분석을 위해서는 자료의 히핑(heaping) 정도를 조사방법에 따라 비교하는 등의 추가적인 과정이 필요할 것이다. 한편, 1차 비임금근로자의 경우 CAPI에서 무응답률이 0%였음에도 불구하고 평균소득 무응답률의 유의한 차이는 발견되지 않았다.

1차 임금근로자의 평균소득은 무응답률에 유의한 차이가 있음에도 불구하고 분포와 평균에는 유의한 차이가 없는 것으로 나타났는데, 이는 정확한 수치를 모르는 응답자가 대략적인 수치라도 대답하는 대신 모름/무응답을 선택하기 때문으로 생각된다.

## 5. 설문유형 변경 실험 분석

### 가. 사회보험 설문항목 방식 변경

개인 설문조사에서 면접원 대면면접의 방식인 CAPI와 개인응답자가 직접 기록하는 웹 전용의 CAWI 집단 간 응답 특성의 차이를 살펴보았다. 개인용 설문에서는 응답자의 사회보험 수급 여부를 조사한다. 1차년도 웹서베이 조사에서 CAPI 집단과 CAWI 집단 간의 실업급여 항목에 대한 응답률에서 차이가 있는 것에 주목하여 2차년도 조사에서는 기본 설문항목의 배열 화면을 변경하여 잘 잡히지 않았던 육아휴직급여, 출산전후휴가급여의 응답률에 차이가 있는지 확인해보고자 하였다.

측정문항은 사회보험 수급 여부에 따른 사회보험 종류이다. [그림 3-2]와 같이 사회보험 종류 세부항목을 위치 변경하여 보여주고 있다.

우선 <표 3-45>는 사회보험 수급 종류별 응답률에 있어 CAPI와 CAWI 집단의 차이가 있는지 살펴보았다. 분석 대상은 사회보험 수급 사실이 있거나 받고 있는 응답자들이다. 1차년도에는 휴업급여, 육아휴직급여, 출산전후휴가급여, 기타 항목에서 CAPI 집단보다 CAWI 집단의 응답률이 통계적으로 유의하게 높은 것으로 보여진다. 2차년도는 노령연금, 육아휴직급여, 출산전후휴가급여, 실업급여, 기타 항목에서 CAWI 집단의 응답률이 통계적으로 유의하게 높은 것으로 나타났다.

<표 3-46>은 사회보험 항목 배열 변경에 따른 차이 여부를 비교하였다. 사회보험 종류의 기존 항목과 항목변경 후의 자료를 묶어 그 차이가 있는지를 분석해 보고자 하였다. 응답률의 차수별 차이를 살펴본 결과 통계적으로 유의한 차이가 없는 것으로 나타났다.

[그림 3-2] 설문항목 방식

화면 구성																															
(기존) 항목	《보기 : 사회보험의 종류》																														
	<table> <tr> <th>국민연금</th><th>특수직역연금</th><th>산재보험</th><th>보훈연금 및 고용보험</th><th>기타</th></tr> <tr> <td>(1) 노령연금</td><td>(6) 사학연금</td><td>(10) 휴업급여</td><td>(13) 보훈연금</td><td>(17) 특례노령연금</td></tr> <tr> <td>(2) 장애연금</td><td>(7) 공무원연금</td><td>(11) 장해연금</td><td>(14) 육아휴직급여</td><td>(18) 기타</td></tr> <tr> <td>(3) 유족연금</td><td>(8) 군인연금</td><td>(12) 유족급여</td><td>(15) 출산전후휴가급여</td><td></td></tr> <tr> <td>(4) 사망일시금</td><td>(9) 반환일시금</td><td></td><td>(16) 실업급여</td><td></td></tr> <tr> <td>(5) 반환일시금</td><td></td><td></td><td></td><td></td></tr> </table>	국민연금	특수직역연금	산재보험	보훈연금 및 고용보험	기타	(1) 노령연금	(6) 사학연금	(10) 휴업급여	(13) 보훈연금	(17) 특례노령연금	(2) 장애연금	(7) 공무원연금	(11) 장해연금	(14) 육아휴직급여	(18) 기타	(3) 유족연금	(8) 군인연금	(12) 유족급여	(15) 출산전후휴가급여		(4) 사망일시금	(9) 반환일시금		(16) 실업급여		(5) 반환일시금				
국민연금	특수직역연금	산재보험	보훈연금 및 고용보험	기타																											
(1) 노령연금	(6) 사학연금	(10) 휴업급여	(13) 보훈연금	(17) 특례노령연금																											
(2) 장애연금	(7) 공무원연금	(11) 장해연금	(14) 육아휴직급여	(18) 기타																											
(3) 유족연금	(8) 군인연금	(12) 유족급여	(15) 출산전후휴가급여																												
(4) 사망일시금	(9) 반환일시금		(16) 실업급여																												
(5) 반환일시금																															
《보기 : 사회보험의 종류》																															
(변경) 항목	<table> <tr> <th>보훈연금 및 고용보험</th><th>산재보험</th><th>국민연금</th><th>특수직역연금</th><th>기타</th></tr> <tr> <td>(1) 보훈연금</td><td>(5) 휴업급여</td><td>(8) 노령연금</td><td>(13) 사학연금</td><td>(17) 특례노령연금</td></tr> <tr> <td>(2) 육아휴직급여</td><td>(6) 장해연금</td><td>(9) 장애연금</td><td>(14) 공무원연금</td><td>(18) 기타</td></tr> <tr> <td>(3) 출산전후휴가급여</td><td>(7) 유족급여</td><td>(10) 유족연금</td><td>(15) 군인연금</td><td></td></tr> <tr> <td>(4) 실업급여</td><td></td><td>(11) 사망일시금</td><td>(16) 반환일시금</td><td></td></tr> <tr> <td></td><td></td><td>(12) 반환일시금</td><td></td><td></td></tr> </table>	보훈연금 및 고용보험	산재보험	국민연금	특수직역연금	기타	(1) 보훈연금	(5) 휴업급여	(8) 노령연금	(13) 사학연금	(17) 특례노령연금	(2) 육아휴직급여	(6) 장해연금	(9) 장애연금	(14) 공무원연금	(18) 기타	(3) 출산전후휴가급여	(7) 유족급여	(10) 유족연금	(15) 군인연금		(4) 실업급여		(11) 사망일시금	(16) 반환일시금				(12) 반환일시금		
보훈연금 및 고용보험	산재보험	국민연금	특수직역연금	기타																											
(1) 보훈연금	(5) 휴업급여	(8) 노령연금	(13) 사학연금	(17) 특례노령연금																											
(2) 육아휴직급여	(6) 장해연금	(9) 장애연금	(14) 공무원연금	(18) 기타																											
(3) 출산전후휴가급여	(7) 유족급여	(10) 유족연금	(15) 군인연금																												
(4) 실업급여		(11) 사망일시금	(16) 반환일시금																												
		(12) 반환일시금																													

〈표 3-45〉 사회보험 수급종류별 응답률의 집단 간 차이 분석

		CAPI		CAWI		diff1)	t-value	Pr> t
		평균	S.D.	평균	S.D.			
1차	노령연금	0.1250	0.3337	0.0543	0.2273	0.0707	1.48	0.1425
	장애연금	0.0000	0.0000	0.0054	0.0737	-0.0054	-1	0.3186
	유족연금	0.0357	0.1873	0.0163	0.1270	0.0194	0.73	0.4699
	사망일시금	0.0000	0.0000	0.0054	0.0737	-0.0054	-1	0.3186
	반환일시금	0.0000	0.0000	0.0163	0.1270	-0.0163	-1.74	0.0833
	사학연금							
	공무원연금	0.0714	0.2599	0.0380	0.1918	0.0334	0.89	0.3761
	군인연금	0.0179	0.1336	0.0054	0.0737	0.0124	0.67	0.5081
	반환일시금	0.0179	0.1336	0.0109	0.1040	0.0070	0.36	0.7202
	휴업급여	0.0000	0.0000	0.0380	0.1918	-0.0380	-2.69	0.0078
	장해연금	0.0179	0.1336	0.0000	0.0000	0.0179	1	0.3217
	유족급여	0.0000	0.0000	0.0054	0.0737	-0.0054	-1	0.3186
	보훈연금	0.0000	0.0000	0.0054	0.0737	-0.0054	-1	0.3186
	육아휴직급여	0.0000	0.0000	0.0217	0.1462	-0.0217	-2.02	0.0452
	출산전후휴가	0.0000	0.0000	0.0272	0.1630	-0.0272	-2.26	0.0249
	실업급여	0.1429	0.3531	0.2500	0.4342	-0.1071	-1.68	0.0935
	특례노령연금							
	기타	0.0179	0.1336	0.0870	0.2825	-0.0691	-2.52	0.0126
2차	노령연금	0.3393	0.4778	0.0326	0.1781	0.3067	4.71	<.0001
	장애연금							
	유족연금	0.0179	0.1336	0.0109	0.1040	0.0070	0.36	0.7202
	사망일시금							
	반환일시금	0.0000	0.0000	0.0163	0.1270	-0.0163	-1.74	0.0833
	사학연금	0.0000	0.0000	0.0163	0.1270	-0.0163	-1.74	0.0833
	공무원연금	0.1071	0.3121	0.0272	0.1630	0.0800	1.84	0.0700
	군인연금	0.0000	0.0000	0.0054	0.0737	-0.0054	-1	0.3186
	반환일시금	0.0000	0.0000	0.0109	0.1040	-0.0109	-1.42	0.1579
	휴업급여	0.0000	0.0000	0.0163	0.1270	-0.0163	-1.74	0.0833
	장해연금	0.0179	0.1336	0.0217	0.1462	-0.0039	-0.18	0.8594
	유족급여	0.0000	0.0000	0.0109	0.1040	-0.0109	-1.42	0.1579
	보훈연금	0.0000	0.0000	0.0109	0.1040	-0.0109	-1.42	0.1579
	육아휴직급여	0.0000	0.0000	0.0272	0.1630	-0.0272	-2.26	0.0249
	출산전후휴가	0.0000	0.0000	0.0217	0.1462	-0.0217	-2.02	0.0452
	실업급여	0.0714	0.2599	0.3043	0.4614	-0.2329	-4.79	<.0001
	특례노령연금	0.0000	0.0000	0.0109	0.104	-0.0109	-1.42	0.1579
	기타	0.0000	0.0000	0.0598	0.2377	-0.0598	-3.41	0.0008

주: 사회보험 수급사실이 있는 응답자 대상.

〈표 3-46〉 사회보험 항목배열 변경에 따른 차이

	1차		2차		diff1)	t-value	Pr> t
	평균	S.D.	평균	S.D.			
노령연금	0.1453	0.3539	0.2033	0.4041	-0.058	-1.18	0.2378
장애연금	0.0086	0.0925	-	-	0.00855	1.00	0.3194
유족연금	0.0427	0.2031	0.0244	0.1549	0.0183	0.78	0.4340
사망일시금	0.0086	0.0925	0.0000	0.0000	0.00855	1.00	0.3194
반환일시금	0.0256	0.1587	0.0244	0.1549	0.00125	0.06	0.9508
사학연금	0.0000	0.0000	0.0244	0.1549	-0.0244	-1.75	0.0833
공무원연금	0.0940	0.2931	0.0894	0.2865	0.00459	0.12	0.9026
군인연금	0.0171	0.1302	0.0081	0.0902	0.00896	0.62	0.5378
반환일시금	0.0256	0.1587	0.0163	0.1270	0.00938	0.50	0.6148
휴업급여	0.0598	0.2382	0.0244	0.1549	0.0354	1.36	0.1757
장해연금	0.0086	0.0925	0.0407	0.1983	-0.0321	-1.62	0.1070
유족급여	0.0086	0.0925	0.0163	0.1270	-0.00771	-0.54	0.5899
보훈연금	0.0086	0.0925	0.0163	0.1270	-0.00771	-0.54	0.5899
육아휴직급여	0.0342	0.1825	0.0407	-0.0065	-0.00646	-0.26	0.7933
출산전후휴가	0.0427	0.2031	0.0325	0.1781	0.0102	0.41	0.6787
실업급여	0.4615	0.5007	0.4878	-0.0263	-0.0263	-0.41	0.6853
특례노령연금	0.0000	0.0000	0.0163	0.1270	-0.0163	-1.42	0.1581
기타	0.1453	0.3539	0.0894	0.2865	0.0559	1.34	0.1816

주: 사회보험 수급사실이 있는 응답자 대상.

## 6. 직산업 코딩 검정

본 항에서는 직산업코드 분포의 조사방법 간 차이를 분석한다. 한국노동패널조사는 응답자들이 종사하는 산업에 관하여 유형 설문에서 “일하는 곳의 주 생산 품목이나 서비스”와 “일하는 곳의 생산품의 원료”를, 직업에 관하여 “주로 하는 일”과 “사용 장비나 도구”, “부서명”을 질문하며, 이를 바탕으로 전문성을 갖춘 코더가 제9차 한국표준산업분류와 제6차 한국표준직업분류에 따른 오픈코드를 지정한다. 한국노동패널조사에서는 직산업코드를 소분류까지 제공하나, 여기서의 분석은 대분류를 이용한다

(김유빈 외, 2017).

직산업코드의 조사방법 간 차이를 분석하기에 앞서 코더 간 불일치율을 살펴볼 필요가 있다. 앞서 설명한 바와 같이, 코더가 수작업으로 오픈 코드를 지정하므로 동일한 응답에 대해서 상이한 코드가 배정될 수 있다. 그러나 자료가 충분히 명확하다면 직산업코드의 코더 간 불일치율은 조사방법에 따라 큰 차이를 보이지 않을 것이다. 본 조사에서는 조사방법이 코더의 의사결정에 영향을 미치지 않도록 2명의 코더에게 조사방법에 대한 정보를 알려주지 않았다(김유빈 외, 2017).

<표 3-47>은 1차와 2차 산업코드의 코더 간 불일치율을 보여준다. 임금 근로자와 비임금근로자 모두 CAWI에서 CAPI보다 불일치율이 높으며 2차 임금근로자를 제외하면 그 차이가 통계적으로 유의함을 알 수 있다. 앞서 언급한 직산업코드 관련 문항들은 모두 응답자가 주관식으로 입력하는데, 적절한 유도가 없는 경우 답변들이 명확하지 못함을 보여주는 결과이다.

<표 3-47> 산업코드의 코더 간 불일치율

		CAPI		CAWI		diff <sup>(1)</sup>	t-value	Pr< t
		불일치율	S.D.	불일치율	S.D.			
1차	임금근로자	0.0366	0.1889	0.1136	0.3183	-0.0771	-2.03	0.04
	비임금근로자	0.0000	0.0000	0.1333	0.3457	-0.1333	-1.72	0.09
2차	임금근로자	0.0864	0.2827	0.0941	0.2929	-0.0077	-0.20	0.84
	비임금근로자	0.0000	0.0000	0.1935	0.4016	-0.1935	-1.98	0.05

주: 1) 산업코드의 코더 간 불일치율의 조사방법에 따른 차이.

<표 3-48> 직업코드의 코더 간 불일치율

		CAPI		CAWI		diff <sup>(1)</sup>	t-value	Pr< t
		불일치율	S.D.	불일치율	S.D.			
1차	임금근로자	0.2317	0.4245	0.2443	0.4309	-0.1261	-0.22	0.83
	비임금근로자	0.1000	0.3078	0.2000	0.4068	-0.1000	-0.93	0.35
2차	임금근로자	0.2099	0.4098	0.2294	0.1656	-0.0195	-0.35	0.73
	비임금근로자	0.0588	0.2425	0.1935	0.4016	-0.1347	-1.26	0.21

주: 1) 직업코드의 코더 간 불일치율의 조사방법에 따른 차이.



<표 3-48>은 직업코드의 코더 간 불일치율을 보여준다. 비임금근로자에 비해 임금근로자의 직업코드 코더 간 불일치율이 높게 나타났으나 각각은 1차와 2차 모두 통계적으로 유의하지 않게 나타났다. 또한 1, 2차 임금근로자와 비임금근로자 모두 CAWI에서 CAPI보다 불일치율이 높게 나타나고 있다.

#### 가. 임금근로자

<표 3-49>는 각 코더별로 지정한 임금근로자의 산업코드 분포의 조사방법 간 차이를 보여준다. 산업코드는 9차 한국표준산업분류상 소분류를 기준으로 조사되었는데 여기서는 9개 주요 산업으로 구분하여 그 분포를 살펴보았다. 1, 2차 조사 모두 응답한 응답자에 한해서 분석하다보니 표본의 한계로 인해 정확한 검정은 불가능했으나 산업코드 분포의 빈도와 비중을 통해 미루어 짐작하였을 때 조사방법 간 차이는 명백하지 않은 것으로 판단된다. 코더와 조사방법에 상관없이 1, 2차 모두 기타서비스업과 광공업이 가장 높은 비중을 차지하고 있는데 이는 공식 고용통계인 통계청의 「경제활동인구조사」의 산업분포와 유사한 결과를 보이고 있다. 반면 농림어업을 제외한 전기·가스 및 수도업과 공공서비스업은 대체로 낮은 비중을 차지한 것으로 나타났다. 두 코더 모두 조사방법별 비중 차이 또한 비중이 가장 높은 광공업과 기타서비스업에서 나타났는데 광공업은 상대적으로 CAPI에서 비중이 낮고 기타서비스업은 CAWI에서 상대적으로 비중이 높게 나타나고 있다.

<표 3-50>은 각 코더가 지정한 임금근로자의 직업코드 분포의 조사방법 간 차이를 나타낸다. 전문관리직과 사무직이 차지하는 비중이 가장 높게 나타나고 있고 나머지 직종들의 비중은 비교적 낮은 것으로 나타났다. 사무직을 제외한 대부분의 직종에서 조사방법 간 비중 차이는 10%p 미만으로 직업코드 분포 역시 조사방법 간 차이가 명백하지 않은 것으로 보인다.

〈표 3-49〉 코더별 임금근로자의 산업코드 분포

		코더1				코더2			
		CAPI		CAWI		CAPI		CAWI	
		빈도	비중	빈도	비중	빈도	비중	빈도	비중
1차	농림어업	0	0.0	0	0.0	0	0.0	0	0.0
	광공업	13	15.9	46	26.7	13	15.9	44	25.1
	건설업	5	6.1	15	8.7	6	7.3	18	10.3
	전기·가스 및 수도업	1	1.2	1	0.6	1	1.2	1	0.6
	도소매 및 음식숙박업	3	3.7	10	5.8	3	3.7	8	4.6
	운수통신업	6	7.3	17	9.9	7	8.5	19	10.9
	금융 및 부동산업	6	7.3	13	7.6	6	7.3	14	8.0
	공공서비스업	8	9.8	9	5.2	8	9.8	9	5.1
	기타서비스업	40	48.8	61	35.5	38	46.3	62	35.4
	전체	82	100.0	172	100.0	82	100.0	175	100.0
2차	농림어업	0	0.0	0	0.0	0	0.0	0	0.0
	광공업	17	21.3	42	24.9	15	19.0	44	26.0
	건설업	5	6.3	17	10.1	5	6.3	16	9.5
	전기·가스 및 수도업	0	0.0	0	0.0	0	0.0	1	0.6
	도소매 및 음식숙박업	4	5.0	12	7.1	5	6.3	11	6.5
	운수통신업	6	7.5	20	11.8	10	12.7	22	13.0
	금융 및 부동산업	6	7.5	13	7.7	5	6.3	12	7.1
	공공서비스업	8	10.0	8	4.7	7	8.9	7	4.1
	기타서비스업	34	42.5	57	33.7	32	40.5	56	33.1
	전체	80	100.0	169	100.0	79	100.0	169	100.0

〈표 3-50〉 코더별 임금근로자의 직업코드 분포

		1차				2차			
		CAPI		CAWI		CAPI		CAWI	
		빈도	비중	빈도	비중	빈도	비중	빈도	비중
1차	전문관리직	37	45.1	69	39.4	37	45.1	72	41.1
	사무직	24	29.3	74	42.3	24	29.3	72	41.1
	서비스직	6	7.3	5	2.9	6	7.3	6	3.4
	농림어업직	0	0.0	0	0.0	0	0.0	0	0.0
	생산직	15	18.3	27	15.4	15	18.3	25	14.3
	전체	82	100.0	175	100.0	82	100.0	175	100.0
2차	전문관리직	33	40.7	66	39.1	37	46.3	64	37.9
	사무직	27	33.3	63	37.3	23	28.8	64	37.9
	서비스직	6	7.4	8	4.7	7	8.8	10	5.9
	농림어업직	0	0.0	0	0.0	0	0.0	1	0.6
	생산직	15	18.5	32	18.9	13	16.3	30	17.8
	전체	81	100.0	169	100.0	80	100.0	169	100.0

#### 나. 비임금근로자

〈표 3-51〉은 각 코더가 지정한 비임금근로자의 산업코드 분포의 조사 방법 간 차이를 보여준다. 비임금근로자의 산업코드 분포의 비중 차이는 임금근로자에 비해서 조사방법에 따라 다소 크게 나타났다. 다만 CAPI로 조사된 표본수는 약 20개, CAWI는 약 30개 남짓으로 임금근로자에 절반에도 못 미치는 수준인데 조사방법에 따른 비중 차이로만 보면 임금근로자에 비해 높으나 빈도 차이는 상대적으로 낮게 나타나고 있어 비임금근로자 역시 조사방법에 따른 산업코드 분포의 차이가 명백하게 나타나지는 않았다. 한편 조사방법별 산업 분포는 기타서비스업과 도소매 및 음식숙박업의 비중이 가장 큰 것은 동일하게 나타났다.

〈표 3-51〉 코더별 비임금근로자의 산업코드 분포

		코더1				코더2			
		CAPI		CAWI		CAPI		CAWI	
		빈도	비중	빈도	비중	빈도	비중	빈도	비중
1차	농림어업	2	10.0	0	0.0	2	10.0	0	0.0
	광공업	2	10.0	3	10.0	2	10.0	6	20.0
	건설업	1	5.0	0	0.0	1	5.0	0	0.0
	전기·가스 및 수도업	0	0.0	0	0.0	0	0.0	0	0.0
	도소매 및 음식숙박업	6	30.0	11	36.7	6	30.0	8	26.7
	운수통신업	0	0.0	4	13.3	0	0.0	4	13.3
	금융 및 부동산업	1	5.0	3	10.0	1	5.0	3	10.0
	공공서비스업	0	0.0	0	0.0	0	0.0	0	0.0
	기타서비스업	8	40.0	9	30.0	8	40.0	9	30.0
	전체	20	100.0	30	100.0	20	100.0	30	100.0
2차	농림어업	1	5.9	1	3.2	1	5.9	1	3.2
	광공업	2	11.8	6	19.4	2	11.8	9	29.0
	건설업	0	0.0	0	0.0	0	0.0	0	0.0
	전기·가스 및 수도업	0	0.0	0	0.0	0	0.0	0	0.0
	도소매 및 음식숙박업	4	23.5	10	32.3	4	23.5	6	19.4
	운수통신업	2	11.8	3	9.7	2	11.8	2	6.5
	금융 및 부동산업	2	11.8	4	12.9	2	11.8	3	9.7
	공공서비스업	0	0.0	0	0.0	0	0.0	0	0.0
	기타서비스업	6	35.3	7	22.6	6	35.3	10	32.3
	전체	17	100.0	31	100.0	17	100.0	31	100.0

〈표 3-52〉는 두 코더가 지정한 비임금근로자의 직업코드 분포의 조사 방법 간 차이를 보여준다. 직업코드 또한 조사방법별 비중의 차이가 10%p 이상인 직종이 있으나 빈도 차이는 5개 미만으로 1, 2차 모두 조사 방법에 따라 직업코드에 차이가 있다고 보기는 어려웠다.

〈표 3-52〉 코더별 비임금근로자의 직업코드 분포

		코더1				코더2			
		CAPI		CAWI		CAPI		CAWI	
		빈도	비중	빈도	비중	빈도	비중	빈도	비중
1차	전문관리직	7	35.0	14	46.7	7	35.0	10	33.3
	사무직	0	0.0	2	6.7	1	5.0	5	16.7
	서비스직	9	45.0	9	30.0	8	40.0	9	30.0
	농림어업직	2	10.0	0	0.0	2	10.0	0	0.0
	생산직	2	10.0	5	16.7	2	10.0	6	20.0
	전체	20	100.0	30	100.0	20	100.0	30	100.0
2차	전문관리직	7	41.2	15	48.4	8	47.1	16	53.3
	사무직	1	5.9	1	3.2	0	0.0	3	10.0
	서비스직	6	35.3	8	25.8	6	35.3	5	16.7
	농림어업직	1	5.9	1	3.2	1	5.9	1	3.3
	생산직	2	11.8	6	19.4	2	11.8	5	16.7
	전체	17	100.0	31	100.0	17	100.0	30	100.0

#### 다. 소 결

본 항에서는 CAPI와 CAWI의 직산업코드의 분포가 조사방법 간 차이가 유의한지 분석하였다. 우선 코더 간 불일치율의 조사방법 간 차이를 살펴본 결과, 코더 간 불일치율은 산업코드와 직업코드 모두 CAWI에서 CAPI보다 높은 것으로 나타났는데 산업코드의 경우 2차 임금근로자를 제외한 대부분의 결과가 조사방법 간 차이가 통계적으로 유의하게 나타나 CAWI를 통해 조사된 자료가 코더들에게 명확한 정보를 주지 못하고 있음을 추론할 수 있다. 반면 직업코드의 경우 코더 간 불일치율 차이가 상대적으로 낮고 통계적으로 유의하지도 않게 나타나 상반된 결과를 보였다. 또한 직산업 코드의 분포를 살펴본 결과 조사방법 간에 뚜렷한 차이를 관찰할 수는 없었다.

다만 산업코드 불일치율에서 상대적으로 CAWI를 통해 조사된 자료가 코더들에게 명확한 정보를 주지 못하는 것으로 보이고 또한 본문에서는

제시되지는 않았지만 직산업 소분류를 기준으로 코더 간 불일치율의 조사 방법 간 차이를 살펴보면 주요 산업으로 한정했을 때에 비해 그 차이가 더 뚜렷한 것으로 나타났다. 향후 CAPI에서 CAWI로 이행함에 있어 오픈코드 지정을 위한 주관식 답변에 대해 적절한 가이드라인 제시가 중요해 보인다.

## 제4절 소 결<sup>21)</sup>

한국노동패널조사는 20년이 넘게 진행되어 오면서, 추적불가나 응답거절 등의 이유로 표본이탈이 발생하여 왔다. 특히 근래에는 젊은 응답자들의 패널 유지를 위해 대면조사의 한계를 극복하고자 웹조사를 대안으로 검토하면서, 자료의 신뢰성 검증의 일환으로 웹조사 실험 연구를 2년에 걸쳐 진행하였다.

1차 웹조사의 목적은 ① 면접원이 없어도 응답자가 설문을 잘 이해하고 응답하는가? ② 조사방법에 따라 설문응답의 일관성이 유지되는가?, 즉 주요 문항들에서 CAPI와 CAWI의 응답값이 차이가 나지 않는가? ③ 조사도구의 변화에도 현재의 조사성공률을 유지할 수 있을까를 살펴보는 것이었다.

2차 CAWI 실험의 목적은 ① 조사도구별 유지율의 차이와 표본 이탈의 편향이 발생하는가?, 즉 유지된 집단과 탈락한 집단이 조사도구별로 다른 특성을 가지고 있는가? ② 2차 웹조사 실험 결과는 1차 조사 결과와 어떠한 차이가 있는가? ③ 1차 실험 결과에 따라 수정된 설문문항의 경우, 2차 웹조사 결과와는 차이가 있는가이다.

이상의 목적에 따라 본 절에서는 1차 웹조사의 결과와 2차 웹조사의 결과를 통합하여 요약 정리하고자 한다. 1차 2차 균형패널 간의 평균값을

21) 본 절은 웹조사 1차 실험과 2차 실험의 결과를 통합하여 정리하기 위해 <표 3-53> ~ <표 3-60>에 ‘비고(1차)’라는 항목을 제시하였는데, 이는 1차 웹조사 결과이다. 따라서 각각 표의 ‘비고(1차)’와 관련하여서는 김유빈 외(2017), 『패널자료 품질개선 연구(VIII)』 제7장 『웹을 활용한 조사의 효과에 대한 실증연구: 노동패널 CAWI 실험 데이터의 분석』(이지은·신선옥·이혜정·김기홍 저술)을 참고하기 바란다.

비교하는 항목은 평균값을 기술해, CAPI와 CAWI 중 높은 항목이 무엇인지 알도록 하였다. 분포를 비교하는 항목은 지면상의 문제로 유의성만 표시하였으므로, 자세한 내용은 본문을 참조하기 바란다. 참고로 1차 조사결과를 비교항목에 표시하였다.

<표 3-53>과 <표 3-54>는 조사도구별로 조사성공과 패널유지에 차이가 있는지 검증한 것이다. CAWI가 CAPI에 비해 진입장벽이 낮아 조사시작은 높지만, 시작 이후 설문이 완성되어야 하는 조사 성공률은 가구와 개인 모두 CAPI가 높았다. 또한 가구원수 대비 응답가구원과 본인응답여부도 CAPI가 높았다. 조사도구별로 응답대상 간의 차이를 분석한 결과, 성별과 학력은 차이가 없었지만, 연령은 CAPI가 CAWI보다 높아 확실히 젊은 연령에서 CAWI가 선호되는 것을 알 수 있다.

패널 유지를 살펴보면, 가구와 개인 모두 패널 유지율에서 조사도구별로 차이가 발생하지 않았다. 성별과 학력은 유의하지 않았고, 연령은 조

<표 3-53> 조사 성공 관련 1차 조사결과 요약

	1차		통계적 유의성
	CAPI	CAWI	
조사시작률	0.6357	0.9235	○
조사성공률(가구)	0.8483	0.6202	○
조사성공률(개인)	0.9877	0.8977	○
전체 표본 대비 조사성공률	0.5393	0.5728	×
대상가구원수 대비 응답가구원수 비율	0.9868	0.8115	○
가구 응답시간(분)	27.37	27.40	×
개인 응답시간(분)	29.45	39.51	○
대리응답(본인여부)	0.973	0.829	○
성별(남성비율)	0.5217	0.5599	×
연령(만나이)	42.64	40.81	□
교육(연수)	14.64	14.45	×

주: 통계적 유의성 ○: 0.01이하, △: 0.05이하, □: 0.1 이하, ×: 유효하지 않음

자료: 김유빈 외(2017), 『패널자료 품질개선 연구(VII)』 제7장 「웹을 활용한 조사의 효과에 대한 실증연구: 노동패널 CAWI 실험 데이터의 분석」(이지은·신선옥·이혜정·김기홍 저술) 재구성.

사도구별 차이가 발생하였는데, CAPI와 CAWI 모두 유지집단보다 탈락 집단이 젊은 연령층이었다. pid 오류는 본 조사과정에서 해결될 수 있는 부분이지만, 오류유형 중 수정이 불가능한 오류가 CAWI에서 많이 발생하여 각별한 주의를 요했다(표 3-54 참조).

<표 3-55>는 액수와 관련된 변수의 응답률 차이이다. 예를 들어 근로소득액 응답률이란 근로소득이 있다고 응답한 비율이다. 1차 균형패널과 2차 균형패널 모두에서 0.05%에서 유의한 차이를 보이는 항목은 가구자료의 부동산소득액, 지난 한 달간 소득뿐이다. 1차와 2차 중 한 개로도 유의한 차이를 보이는 항목은 소득총액, 이전소득, 금융자산, 월평균저축액이다. 즉 액수 관련 변수의 응답률은 한 차수만을 봤을 때는 많은 항목들이 유의한 차이를 보이지만, 2개 차수를 모두 고려하면, 그 유의성이 낮아져, CAPI와 CAWI에 확인한 차이가 있다고 확인할 수는 없었다.

<표 3-56>은 액수 관련 변수에 대한 평균값의 비교이다. 응답률에서 일부 항목이 조사도구별 차이를 보인 반면, 평균값에서는 차이를 보인 항목이 하나도 없었다.

<표 3-54> 2차 패널 유지 결과 요약

		2차		통계적 유의성
		CAPI	CAWI	
패널유지율	가구	0.6291	0.6775	×
	개인	0.05490	0.5481	×
성별(남성비율)	유지	52.9	59.5	×
	탈락	51.3	50.9	
연령(만나이)	유지	44.81	42.08	○
	탈락	41.38	39.37	
학력(중졸, 고졸, 전문대졸, 대졸, 석사 이상)		-	-	×
지역(서울권, 경인권, 충청권, 영남권, 호남권)		-	-	△
pid 오류율		0.1232	0.0983	×
pid 오류 유형(가구원내 가구원번호 오류, 가구원추가시 가구원번호 오류, 이름 오류, 가구 전체 모두 다른 이름, 기타)		-	-	○

주: 통계적 유의성 ○: 0.01이하, △: 0.05이하, □: 0.1 이하, ×: 유효하지 않음.



〈표 3-55〉 1차 2차 균형패널 조사결과 요약(액수 관련 변수 응답률)

		1차균형패널		통계적 유의성	2차균형패널		통계적 유의성	비고 (1차)
		CAPI	CAWI		CAPI	CAWI		
가구	소득총액	0.9895	0.9471	△	0.9789	0.9423	□	×
	근로소득액	0.9158	0.8462	□	0.9368	0.8894	×	□
	금융소득액	0.4316	0.4615	×	0.4737	0.4567	×	△
	부동산소득액	0.1895	0.0865	△	0.1684	0.0721	△	△
	이전소득액	0.5053	0.2308	○	0.1158	0.0529	□	○
	기타소득액	0.3368	0.3077	×	0.2526	0.2837	×	×
	금융자산	0.9684	0.8606	○	1.0	1.0	-	○
	지난 한달간 소득	0.9789	0.8846	○	0.9684	0.8846	○	○
	월평균저축액	0.9053	0.8269	□	0.9053	0.8029	△	□
	생활비	1.0	1.0	-	1.0	0.9856	□	×
개인 평균 소득	임금근로자	1.0	0.989	×	1.0	1.0	×	×
	비임금근로자	0.95	0.793	×	1.0	0.84.4	□	×

주: 통계적 유의성 ○: 0.01이하, △: 0.05이하, □: 0.1 이하, ×: 유효하지 않음.

〈표 3-56〉 1차 2차 균형패널 조사결과 요약(액수 관련 변수 분포- 평균 비교)  
(단위: 만 원)

		1차균형패널		통계적 유의성	2차균형패널		통계적 유의성	비고 (1차)
		CAPI	CAWI		CAPI	CAWI		
가구	근로소득액	5,229.4	4,964.7	×	5,292.8	4,870.8	×	×
	금융소득액	329.2	756.4	×	628.5	648.5	×	×
	부동산소득액	1,342.4	2,155.3	△	3,306.7	1,844.3	×	×
	이전소득액	2,279.2	1,737.8	×	333.2	205.6	×	×
	기타소득액	1,376.7	1,343.3	×	1,354.9	1,281.8	×	×
	금융자산	6,257.1	6,788.1	×	7,880.9	6,773.3	×	×
	지난 한 달 소득	432.7	387.1	×	467.9	418.8	×	×
	월평균 저축액	109.7	88.2	×	110.3	103.4	×	×
개인 평균 소득	임금근로자	284.2	292.8	×	324.7	313.9	×	×
	비임금근로자	279.5	273.8	×	247.8	296.8	×	×

주: 통계적 유의성 ○: 0.01이하, △: 0.05이하, □: 0.1 이하, ×: 유효하지 않음.

<표 3-57>은 일자리 관련 변수와 직업훈련 관련 변수에서 조사도구별 비중 차이를 나타낸 것이다. 평균 자격증 보유개수만 통계적으로 유의한 차이를 보이며 그 외에는 차이가 없는 것으로 나타났다.

<표 3-58>은 개인의 성향이나 개인 정보 등과 같이 응답자가 응답을 꺼릴 수 있는 민감문항의 조사도구별 응답률 차이를 살펴본 것이다. 1차 2차 5% 수준에서 통계적으로 유의한 차이를 보이는 항목은 흡연여부, 금연 시도 여부, 사회보험 수급여부, 수급 사회보험 개수이다. 금주 시도 여

<표 3-57> 1차 2차 균형패널 조사결과 요약(일자리, 직업훈련)

		1차균형패널		통계적 유의성	2차균형패널		통계적 유의성	비고 (1차)
		CAPI	CAWI		CAPI	CAWI		
임금근로자 비중		0.8039	0.8544	×	0.8182	0.8424	×	×
종업원수	임금	0.7927	0.8011	×	0.8148	0.8012	×	×
	비임금	1.0000	0.9333	×	1.0000	0.9688	×	×
직업훈련 경험 여부		0.1255	0.2554	○	0.1516	0.2021	□	○
직업훈련 받은 개수(개)		1.8500	1.6447	×	1.7895	1.7429	×	-
자격증 보유 여부		0.4902	0.4401	×	0.3755	0.4615	△	×
자격증 보유 개수(개)		1.6267	2.2016	○	1.7018	2.2656	△	-

주: 통계적 유의성 ○: 0.01이하, △: 0.05이하, □: 0.1 이하, ×: 유효하지 않음.

<표 3-58> 1차 2차 균형패널 조사결과 요약(민감문항 응답률)

		1차균형패널		통계적 유의성	2차균형패널		통계적 유의성	비고 (1차)
		CAPI	CAWI		CAPI	CAWI		
흡연 여부		0.1000	0.2043	○	0.0929	0.2222	○	△
금연 시도		0.0286	0.0932	○	0.0357	0.0968	△	○
음주 여부		0.6643	0.6595	×	0.6429	0.6703	×	×
금주 시도		0.0214	0.0323	×	0.0071	0.0609	○	△
사회보험 수급 여부		0.0980	0.1807	○	0.1119	0.1646	△	○
수급사회보험 개수		1.0000	1.4821	○	1.0000	1.4000	○	-
혼인상태 변화 여부		0.0627	0.0806	×	0.0325	0.0304	×	×
혼인상태 변화 횟수		1.1000	1.6522	○	1.7500	1.5385	×	-

주: 통계적 유의성 ○: 0.01이하, △: 0.05이하, □: 0.1 이하, ×: 유효하지 않음.

부는 1차 균형패널에서는 통계적으로 유의하지 않았지만, 2차 균형패널에서는 유의했으며, 1차 응답자만으로 분석한 1차 조사에서도 유의했었다. 즉 응답자가 응답하기를 꺼리는 민감문항은 혼인을 제외한 대부분의

〈표 3-59〉 1차 2차 균형패널 조사결과 요약(직장만족도 응답 차이)

	1차균형패널		통계적 유의성	2차균형패널		통계적 유의성	비고 (1차)
	CAPI	CAWI		CAPI	CAWI		
조직몰입도							
지금 직장은 다닐 만하다	3.6707	3.4716	×	3.6543	3.4269	△	△
이 직장에 들어온 것을 기쁘게 생각한다	3.6463	3.3920	△	3.6420	3.3509	△	○
친구에게 이 직장을 추천하고 싶다	3.4390	3.1023	△	3.5185	3.1228	○	○
이 직장을 다른 사람들에게 자랑할 수 있다	3.5122	3.1932	△	3.4815	3.2281	△	○
이 직장을 계속 다니고 싶다	3.7439	3.5455	×	3.8025	3.5380	△	△
직무만족도							
현재 하고 있는 일에 만족하고 있다	3.6765	3.3058	○	3.6061	3.3350	○	○
이 일을 열정적으로 하고 있다	3.6863	3.3301	○	3.5859	3.3695	△	○
이 일을 즐겁게 하고 있다	3.7059	3.2330	○	3.5859	3.2857	○	○
이 일을 보람을 느끼면서 하고 있다	3.7451	3.2621	○	3.5960	3.2611	○	○
이 일을 계속하고 싶다	3.8529	3.4272	○	3.6667	3.4877	×	○
요인별 직무만족도							
임금/보수	2.9314	3.0777	×	2.7980	2.9901	□	○
취업안정성	2.5000	2.7816	△	2.4747	2.7241	△	○
일의 내용	2.3922	2.7039	○	2.3535	2.6108	○	○
근무환경	2.4216	2.7427	○	2.4848	2.7635	△	○
근로시간	2.3922	2.7282	○	2.5152	2.6700	×	○
개인발전가능성	2.5588	2.8932	○	2.5657	2.9606	○	○
의사소통 및 인간관계	2.3627	2.7573	○	2.4747	2.7438	△	○
인사고과의 공정성	2.7805	3.0057	□	2.6914	3.0643	○	○
복지후생제도	2.9512	3.2159	□	2.8519	3.1980	△	○
전반적 일자리만족도	2.4902	2.8058	○	2.4545	2.7931	○	○
전반적 일반만족도	2.3922	2.7864	○	2.3939	2.7192	○	○

주: 통계적 유의성 ○: 0.01이하, △: 0.05이하, □: 0.1 이하, ×: 유효하지 않음.

문항에서 차이를 보여, CAWI 도입 시 유의해야 한다.

<표 3-59>는 5점 척도인 조직몰입도, 직무만족도, 요인별 직무 만족도, 전반적 일자리 만족도, 전반적 일만족도의 조사도구별 차이를 살펴본 것이다. 대부분의 항목에서 CAPI와 CAWI의 평균값의 차이가 통계적으로 유의한 것으로 나타나 CAWI 도입 시 유의해야 할 것이다.

<표 3-60>은 5점 척도인 생활만족도를 살펴본 것으로, 대부분의 항목에서 CAPI와 CAWI의 평균값의 차이가 통계적으로 유의한 것으로 나타났다.

<표 3-60> 1차 2차 균형패널 조사결과 요약(생활만족도)

	1차균형패널		통계적 유의성	2차균형패널		통계적 유의성	비고 (1차)
	CAPI	CAWI		CAPI	CAWI		
가족수입	2.9071	3.1111	△	2.5214	2.8172	○	○
여가생활	2.8286	3.0968	○	2.8714	3.1613	○	○
주거환경	2.5929	2.9391	○	2.8071	3.1147	○	○
가족관계	2.2786	2.4803	△	2.6429	2.9140	○	○
친인척관계	2.5857	2.8602	○	2.2571	2.4910	○	○
사회적 친분	2.4143	2.7025	○	2.5071	2.8280	○	○
전반적 만족도	2.5357	2.8029	○	2.4214	2.7419	○	○

주: 통계적 유의성 ○: 0.01이하, △: 0.05이하, □: 0.1 이하, ×: 유효하지 않음.

<표 3-61>은 5점 척도로 구성된 건강상태 및 장애, 육체적 제약 등의 항목을 살펴본 것이다. 1차 2차 균형패널 모두에서 5% 수준에서 유의한 항목은 현재 건강상태, 보통사람 대비 건강상태, 육체적 제약, 활동제약3(쇼핑하기, 병원가기 등) 항목이다. 이 외의 항목들은 2차 균형패널에서는 유의하지만, 1차 균형패널에서는 유의하지 않게 나타나 항목별로 상이한 결과를 보였다.

<표 3-62>는 종사상 지위, 비정규직 여부, 근로시간과 관련된 문항의 조사도구별 차이를 살펴보았다. 종사상 지위인 임금근로자 유형과 비임금근로자 유형 모두 통계적 유의한 차이가 없는 것으로 나타났다. 비정규직 여부는 통계적 기준에 의한 구분이 큰 차이를 보였다. 이는 ‘근로시간

〈표 3-61〉 1차 2차 균형패널 조사결과 요약(건강상태 및 장애)

	1차균형패널		통계적 유의성	2차균형패널		통계적 유의성	비고 (1차)
	CAPI	CAWI		CAPI	CAWI		
현재 건강상태	2.4071	2.6918	○	2.4286	2.7706	○	○
1년 전 대비 건강상태	3.0071	3.0645	×	2.9857	3.0824	□	○
보통사람 대비 건강상태	2.6571	2.8495	○	2.6571	2.8746	○	×
감각기관장애	2.0000	1.9391	○	1.9929	1.9677	□	○
육체적 제약	1.9786	1.8961	○	1.9571	1.9032	△	○
활동제약1 (배우기, 기억하기 등)	1.9857	1.9068	○	1.9786	1.9355	△	○
활동제약2 (옷입기, 목욕하기 등)	2.0000	1.9749	○	1.9857	1.9821	×	×
활동제약3 (쇼핑하기, 병원가기 등)	2.0000	1.9677	○	1.9857	1.9534	△	×
활동제약4 (작업활동 어려움)	1.9571	1.9068	△	1.9571	1.9176	□	○

주: 통계적 유의성 ○: 0.01이하, △: 0.05이하, □: 0.1 이하, ×: 유효하지 않음.

〈표 3-62〉 1차 2차 균형패널 조사결과 요약(일자리)

		1차균형패널		통계적 유의성	2차균형패널		통계적 유의성	비고 (1차)
		CAPI	CAWI		CAPI	CAWI		
임금근로자 유형 (상용, 임시, 일용)		-	-	×	-	-	×	×
비임금근로자 유형 (고용주, 자영업자, 무급가족종사자)		-	-	×	-	-	×	×
비정규직 여부	통계기준	0.3659	0.6364	○	0.3457	0.6550	○	○
	자기선언기준	0.2099	0.1257	□	0.1875	0.1412	×	□
	불일치여부 <sup>2)</sup>	0.1975	0.5314	○	0.1750	0.5118	○	○
근로시간 (시간)	임금근로자	42.6	45.9	△	43.3	43.0	×	×
	비임금근로자	42.3	44.2	×	43.7	43.8	×	×

주: 통계적 유의성 ○: 0.01이하, △: 0.05이하, □: 0.1 이하, ×: 유효하지 않음.

의 정함'이라는 문구의 정의를 CAWI 응답자가 인지하지 못했기 때문이라 생각되어, 2차 조사에서 부가적 설명을 넣었음에도 그 차이가 해소되

지 못해, CAWI 도입 시 응답자의 문항의 이해도를 높이는 다양한 방법이 필요한 것으로 보인다. 근로시간의 차이는 없었다.

<표 3-63>은 개방형 질문의 정보를 이용하여 전문가가 수작업으로 숫자를 지정하는 직산업 코딩 불일치율을 살펴보았다. 불일치율은 모든 항목에서 CAPI보다 CAWI에서 더 높아, CAWI가 개방형 질문에 취약함을 암시했다. 1차 조사만을 살펴본 비교항목을 보면, 대부분의 항목에서 유의한 차이를 보였지만, 1차 2차 균형패널로 살펴보았을 때는 임금근로자의 산업코드만 5% 수준에서 통계적으로 유의한 차이를 보였다.

<표 3-63> 1차 2차 균형패널 조사결과 요약(직산업 코딩 불일치율)

		1차균형패널		통계적 유의성	2차균형패널		통계적 유의성	비교 (1차)
		CAPI	CAWI		CAPI	CAWI		
산업	임금근로자	0.0366	0.1136	△	0.0864	0.0941	×	○
	비임금근로자	0.0000	0.1333	□	0.0000	0.1935	□	△
직업	임금근로자	0.2317	0.2443	×	0.2099	0.2294	×	×
	비임금근로자	0.1000	0.2000	×	0.0588	0.1935	×	△

주: 통계적 유의성 ○: 0.01이하, △: 0.05이하, □: 0.1 이하, ×: 유효하지 않음.

이상의 결과를 요약하면 다음과 같다. 우선 패널유지율 면에서는 CAPI와 CAWI에 유의한 차이가 발생하지 않아, CAWI 도입에 있어 긍정적이라 할 수 있다. 다만 CAWI가 조사대상 가구원 대비 응답 가구원수가 적은 점, 응답자의 낮은 연령, 높은 대리응답은 CAWI 도입 시에 고려해야 할 중요한 결과이다.

항목별로는 금액 관련 변수의 응답률에서는 이전소득, 금융자산, 지난한 달간의 소득항목에서 차이가 나타났지만, 각 항목의 평균값에서는 차이가 발생하지 않아 고무적이다. 일자리 관련 문항 중 임금근로자 여부, 종사상 지위, 근로시간, 직산업 코딩 등에는 유의한 차이가 거의 없었으나, 문항의 이해도 차이로 인해 발생한 비정규직 여부의 유의한 차이는 CAWI 도입 시 해결해야 할 과제로 보인다.

민감문항인 흡연, 음주, 사회보험수급여부, 혼인상태 변화여부 중 흡연과 음주, 사회보험수급여부는 통계적으로 유의한 차이를 보였다.

특히 문제가 되는 항목은 5점 척도 문항으로, 조직 몰입도, 직무만족도, 요인별 직무만족도, 전반적 일자리 만족도, 전반적 일 만족도, 생활만족도 등은 모두 유의한 차이가 발생했다. 따라서 CAWI 도입 시 본 문항을 제외할 것을 조심스럽게 제안한다.

신뢰성 있는 자료의 생산을 위해 표본 이탈에 대처하기 위한 다양한 조사방법의 도입은 이제 피할 수 없다. 그러나 표본 유지에만 초점을 맞춰 조사도구의 차이에서 발생할 수 있는 차이를 간과해서는 안 된다. 따라서 노동패널팀에서 실시한 1차, 2차 웹조사 실험연구가 CAWI 도입에 있어 중요한 자료로 사용되길 바란다.

## 제 4 장

## 출산력변수 형성방법 및 문제점 진단

## 제1절 서론

출산력 관련 변수는 주요 인구통계학적 변수로서 중요한 의미가 있다. 특히 우리나라 저출산 문제가 매우 심각해짐에 따라 인구학 연구들이 활발히 진행되고 있어, 출산 관련 변수의 중요성은 더욱 높아지고 있다. 예를 들면, 여성근로자들의 경력단절 및 생애주기 임금을 결정함에 있어 자녀의 출생이 어떠한 영향을 미치는지 연구하는 데 주요변수로서 활용될 수가 있다. 또 최근 활발하게 논의 되고 있는 일·가정 양립정책과 관련된 연구들에서 남성과 여성의 자녀수 변수를 필요로 하는 경우가 많은데, 이를 위해서도 반드시 필요한 변수이다.

노동패널조사에서는 개인이 패널에 처음 진입할 당시 (신규조사자) 출산이력에 관하여 직접적으로 질문하지만 그 이후 차수에서는 직접적으로 출산이력 변화에 대한 질문을 하지 않는다. 따라서 정확한 출산이력의 변화를 알아내기 위해서는 가구조사에서 사교육 관련 문항, 가구원 정보 등 다양한 변수들을 활용하여 간접적으로 파악하여야 한다. 이때, 변화하는 자녀수 이력을 구축하는 과정이 매우 복잡하며 자녀수 변수를 연구자 필요에 따라 정의하고 가공하여야 하는 문제가 발생한다. 따라서 자료 활용의 편의성이 떨어지는 문제점과 연구자가 자의적으로 변수를 정의하고



생성하여야 해서 같은 노동패널을 활용한 연구 간에도 출산변수가 상이하게 정의됨은 물론 표본수나 분포가 달라질 수 있다.

특히 노동패널은 조사표본을 장기간에 걸쳐 추적할 수 있어 패널모형을 활용하는 데 강점을 갖지만, 출산력 관련 문항을 이용하고자 조사 차수를 연결하여 추적하는 과정에서 연결 정보의 정확도에 따라 상당수 표본이 누락될 우려가 있다. 그 과정이 매우 복잡하여 출산 관련 연구를 하는 연구자에게는 한계점으로 작용한다.

따라서 출산력 자료 구축 시 활용할 수 있는 변수들과 구축방법들을 정리하여 연구자의 출산력변수 이해를 증진시키고 자료처리과정을 단축시키기 위하여 본 연구를 진행하였다. 또한 출산력변수 생성과정에서 발생하는 여러 문제점을 논의하여 노동패널 출산력 데이터의 품질개선을 위한 방안을 제시한다.

## 제2절 출산력 관련 설문문항 및 변수 검토

노동패널조사에서 출산력 정보를 구축하기 위하여 사용할 수 있는 문항은 크게 출산 관련 문항과 사교육 문항이 있다. 출산 관련 문항은 신규패널 진입자에게 초산년도와 자녀수를 묻고있으며, 그 후 자녀수 변화에 대해 직접 묻지는 않는다. 사교육 문항은 매년 사교육이 필요한 0세~대학생(대학원생) 자녀의 가구원번호와 엄마의 가구원번호를 묻고 있어 사교육 문항을 활용하여 자녀에 대한 정보를 엄마 정보와 연결할 수 있다. 따라서 신규진입 당시 응답한 출산경험 문항과 사교육 문항을 활용하여 출산이력변수-미취학 자녀의 수, 출산시기, 해당 자녀 엄마의 가구원번호 및 나이 등-를 구축할 수 있다. 그런데 가구조사와 개인조사를 연결하는 과정, 패널로 구축하는 과정, 기본적인 변수를 생성하는 과정에서 몇 가지 문제점이 발생할 수 있어 변수활용에 어려움이 발생한다. 따라서 본 절에서는 변수생성과정들에 관한 논의를 하기 이전에 활용 가능한 변수별 특성과 특이점에 대하여 논의한다.

출산력 정보를 구축하기 위하여 사용할 수 있는 변수들은 다음의 <표 4-1>에 정리된 바와 같다. 먼저, 신규조사자가 패널에 처음 진입하였을 때 출산 이력에 관하여 질문(P\_9071~75)한다. 구체적으로 자녀 출산경험이 있는지 묻고, 출산경험이 있다고 응답한 자에 대해 총 자녀수를 성별로 묻고(아들, 딸) 초산시기를 묻는다. 이때, 자녀의 연령이나 생년월일은 직접 묻지 않고 있다. 따라서 이 변수를 염두해 두고 연구자가 노동패널로 특정 분석을 하려고 할 때, 실제 사용상 여러 가지 어려움이 있을 수 있다.<sup>22)</sup>

첫 번째 문제점은 신규조사자에게만 질문하는 변수이므로, 이 변수만으로는 패널 데이터로 구축하는 것이 거의 불가능하다. 만약 출산이력 관련 변수들을 패널로 이어붙인 후 그대로 사용하려고 한다면, 이 변수는 패널 진입 차수 후로는 불관측(missing) 처리되어 있어 사용할 수 없다. 다만, 패널 진입 당시 이미 가임기가 지난 연령(만 45세 이상)을 대상으로 하여도 되는 분석을 할 때는, 이 변수만을 활용해도 패널로 사용할 수는 있다(불관측 차수에 대해 그전 차수의 정보로 대체하는 방식으로). 또한, 간단한 분석을 위해서라면 노동패널 신규응답자들만을 통합(pooling)하여 통합횡단면(pooled cross-section)의 형태로 분석할 수 있다. 학생 연구자나 데이터 가공이 익숙하지 않은 연구자들은 이 방법을 활용하면 간단하게 출산변수를 사용할 수 있다. 단, 이 변수는 1차년도에는 응답 순서가 반대여서 데이터를 활용할 때 1차년수의 코드값을 변경해 주어야 한다는 점에 유의해야 한다(2차년도 이후에는 출산하였음이 “2”로 코딩되어 있으나 1차년도에는 “1”로 코딩되어 있다).

두 번째는, 출산이력 관련 질문을 활용하여 남성에 대해 분석을 할 때는 설문이 일관적으로 진행되지 않아 변수 활용에 유의하여야 한다. 남성의 경우, 1~10차년도까지 출산 관련 문항이 질문되었으나, 11차년도 이후에는 질문되지 않은 것으로 파악된다. 따라서 <표 4-2>에서 정리된 바와 같이, 신규응답자의 아들, 딸 수의 관측치가 11차년도 이후 남성에게서 사라지는 것이 확인되며, 통합횡단면(pooled-cross-section)으로

22) 예컨대 초산에 사망한 자녀가 있을 시 실제 자녀수와 초산시기 정보가 관련이 없는 정보일 수 있다.

〈표 4-1〉 출산력 정보 구축을 위해 사용할 수 있는 변수

설문문항 유형	변수명	질문내용 및 조사 차수
출산 관련 문항	p_9071: 출산경험 유무	- 1차부터 신규조사자만 조사
	p_9075: 자녀 초산시기 (년)	
	p_9073: 자녀 초산시기 (월)	
	p_9074: 자녀수 (아들)	
	p_9075: 자녀수 (딸)	
사교육 관련 문항	h_1501: 0세~고등학생(재수생) 이하 자녀 유무	- 1차, 3~19차에만 조사 - “_____님 댁에는 0세 이상 고등학생 이하(재수생 포함)의 자녀가 있습니까?”
	h_1502: 0세~고등학생(재수생) 이하 자녀수	
	h_1511, h1611, h1711, h1811, h1911: 자녀 1~5의 가구원번호	
	h_1512, h_1612, h_1712, h_1812, h_1912: 자녀1~5의 엄마의 가구원번호	- 16차부터 미취학자녀(0~5세)와 0~고등학생(재수생) 이하 자녀수를 분리하여 질문함 - “_____님 댁에는 0세 이상 고등학생 이하(재수생 포함)의 자녀가 있습니까?” - “고등학생 이하 자녀 중 만 0~5세 이하 자녀가 있습니까?”
	h_1503: 미취학자녀 유무	
	h_1504: 취학-재수생 자녀 유무	
	h_1505: 미취학 자녀 수	- 7차부터 조사 - “___님 댁에서는 대학(원)생인 자녀가 있습니까? (휴학 중인 자녀도 포함됩니다.)”
	h_1506: 고등학생(재수생) 이하 자녀 수	
	h_2001: 대학생 이상 자녀 유무	(직접 질문하는 문항은 없음)
	h_2002: 대학생 이상 자녀 수	자녀 일련 번호
부모와의 교류-사적 이전문항 <sup>23)</sup>	h_2011, h_2021, h_2031, h_2041, h_2051: 대학생 이상 자녀의 가구원번호	엄마 일련 번호
	h_2012, h_2022, h_2032, h_2042, h_2052: 대학생 이상 자녀의 엄마의 가구원번호	- 9차부터 조사 - “※ 다음은 따로 사는 자녀에 대한 질문 입니다. 장성하여 경제적으로 독립해 따 로 살고 있는 가구주의 자녀가 있는 경 우에만 응답해 주십시오. 가구주의 자녀 중 작년(2012년) 한 해 동안 한 달 이상 따로 살고 있었던 자녀가 있습니까?” “따로 살고 있었던 자녀는 몇이었습니까?”
	h_1301: 따로 살고 있던 자녀 여부	
	h_1302: 따로 살고 있던 자녀 수	

〈표 4-2〉 출산력 신규 응답자의 자녀수 문항 남녀 비교( $p_{9074}+p_{9075}$ )

차수	남성 기혼 자녀수 응답자	여성 기혼 자녀수 응답자	자녀수 평균 (남성)	자녀수 평균 (여성)
1	4,201	4,890	2.52	2.70
2	14	11	1.86	3.00
3	7	4	2.57	2.25
4	16	16	1.38	1.44
5	7	8	1.00	1.38
6	6	10	1.83	1.50
7	5	11	1.00	1.73
8	5	7	1.00	2.14
9	8	7	1.63	1.29
10	3	8	2.00	1.25
11		3		1.33
12		1,152		2.89
13		8		1.13
14		5		1.00
15		4		1.25
16		2		1.50
17		3		1.67
18		4		1.25
19		3		1.00

이 변수를 활용하게 되면 때 남성의 관측치가 확연하게 감소한다.

따라서 이 변수를 활용한 단순분석을 하고자 할 때는, 여성만을 표본으로 한정된 분석만 가능하며, 남성을 활용하고자 할 때는 가구주와의 관계( $p_{0102}$ ) 변수를 활용하여 부부를 매칭시킨 후 남성에게 아내의 자녀수를 이어 붙여야만 사용할 수 있다. 이 때, 가구주와의 관계변수를 활용하여 부부를 매칭시키는 과정은 매우 복잡하여 이용자의 편의성이 떨어지는 문제가 발생한다.

23) 이 변수들은 자녀가 장성한 연령대(60대 이상)의 표본들의 총 자녀수를 집계할 때 사용될 수 있다.

다음으로 사용할 수 있는 변수들은 사교육 관련 문항들이다. 사교육 관련 문항들은 구체적으로 0세~고등학생(재수생) 이하 자녀의 수를 묻고 있으며, 각 자녀별로 자녀의 가구원번호와 자녀 엄마의 가구원번호를 묻고 있어 연구목적에 따라 매년 추적되는 자녀수 변수를 그대로 사용할 수 있다는 장점이 있다. 예를 들어 0세~고등학생(재수생) 이하 자녀 유무와 자녀수에만 초점을 맞추게 된다면, 가구단위의 분석일 경우 그대로 가구 단위 패널로 데이터를 수직 연결하여 사용할 수 있다. 단, 개인 단위의 분석을 할 때 이 변수를 그대로 사용하려면 엄마의 정보를 연결해야 되는데, 가구에서 발견되는 여러 자녀의 엄마가 다른 경우가 있을 경우[예를 들어 3대가 함께 거주할 때 조카와 삼촌이 모두 고등학생(재수생) 이하 자녀일 경우], 자녀수 연결이 복잡해진다. 따라서 연구자의 자의적인 판단에 의해 표본을 선정하고 분석을 제한하는 것이 필요할 수 있다.

특히 사교육 문항을 사용하였을 때 주의할 점이 있다. 아빠에게 자녀를 연결할 때는 가구주와의 관계변수를 활용하여 엄마의 배우자를 찾아낸 후 추정된 엄마의 자녀수를 배우자에게 연결해 주어야 하는 어려움이 있다는 점이다. 또한, 자녀들의 엄마가 가구에 함께 살고 있지 않을 때는 가구에 있는 자녀가 엄마와 연결되지 못하는 문제가 생길 수 있다. 만약 아빠와 거주하는 자녀들인지를 나타내려면 가구주와의 관계 변수로 판단하고 아빠와 연결해 주어야 한다. 그런데 가구주와의 관계 변수는 케이스가 워낙 다양해서 일부 코딩에러가 있는 것으로 보이며, 이를 활용해 생성하는 모든 자녀수 변수들은 일정 부분 오류를 포함하게 된다.<sup>24)</sup>

사교육 문항 사용 시 또 다른 문제점은, 사교육 문항에서 응답되는 자녀의 엄마가 재혼가정 등의 이유로 친모가 아닐 수 있어, 자녀수를 엄마에게 연결하였을 시, 해당 엄마의 실제 출산 이력이 오측될 수 있다. 이러한 경우는 드물 수 있지만, 실제 패널로 출산이력을 구축하게 되면 구축을 위해 사용되는 변수들이 다양하고, 많고, 매 차수마다 미세하게 차이

24) 예컨대, 가구원이 출산으로 추가된 경우, 해당 가구원(출생자녀)의 가구원 코드로 가구주나 가구주의 배우자 등 논리적으로 맞지 않는 코드가 부여되는 경우가 간혹 있다. 이것은 가구원 추가 탈락 이유를 잘못 코딩한 것일 수도 있고, 가구주와의 관계 부분을 잘못 코딩한 것일 수도 있지만, 아무쪼록 가구주와의 관계변수에서 코딩에러가 일부 있을 가능성이 존재함은 분명하다.

점이 있을 수가 있어, 너무나 다양하게 고려해야 할 특이 케이스들이 발생하여, 이 한가지의 케이스도 연구자에게 데이터 사용의 어려움으로 작용한다.

한편, 여성의 출산과 경력단절 등의 연구에서 흔히 사용되는 미취학 자녀수 변수는 16차 조사부터 0~고교생(재학생) 이하 자녀수 변수와 별도로 조사되어 왔다. 따라서 연구자가 미취학 자녀수 변수가 필요하다면, 16차 조사 이후 조사들을 사용하여 미취학 자녀수 변수를 활용할 수 있다. 다만 주의할 점은, 미취학 자녀수 변수를 생성하기 위해 노동패널 조사는 “고등학생 이하 자녀 중 만 0~5세 이하 자녀가 있습니까?” 라고 질문하기 때문에, 이 때 응답된 자녀수가 실제 미취학 자녀수와 차이가 날 수 있다(표 4-3 참조).

또한, 출산이력 패널 데이터 구축을 위하여 사교육 문항을 사용하게 될 때, 이 문항은 2차년도에서 조사되지 않았기 때문에 노동패널의 장점인 긴 패널조사기간(1~19차)을 모두 사용하지 못하는 표본이 다수 발생할 수 있어 주의하여야 한다. 1차에서만 조사되고 3차 조사부터 표본에서 탈락한 경우 2차년도의 자녀수에 대해서는 알 수 없다.<sup>25)</sup> 이 표본들의 경우 2차년도의 가구원 구성과 가구주와의 관계 변수를 사용하여 자녀수를 impute 해주어야 사용할 수 있는데, 앞서 언급 하였듯 이 과정이 상당히 복잡하여 데이터 접근성이 떨어진다. 만약 부부가 1차년도와 2차년도 사이에 이혼이나 별거를 하면서 엄마가 자녀를 데리고 떨어져 살게 되는 등

〈표 4-3〉 2016년도 미취학아동 가구원수 비교

0~5세 이하 자녀 수	연령(5세 미만) 이용 자녀수	사교육 문항 “미취학자녀 수” 응답 케이스 수
1	607	648
2	197	281
3	11	16
계	815	945

주: 0명 응답자는 제외되었음.

25) 개인조사를 기준으로: 789명(1차, 2차 개인조사 있고, 3차 없는 경우 489명; 1차 개인조사 있고 2, 3차 없고 바로 4차인 경우 300명).

가구 형태에 변화가 생겼다면 아버지 쪽에서는 자녀가 실제로 존재함에도 불구하고 자녀정보가 누락될 수 있다.<sup>26)</sup> 따라서 특이한 관측치들이 발생하게 되고 이러한 관측치에 대하여 연구자가 자의적으로 변수에 대한 정의를 변경하거나 연구 대상을 제한하여야 하는 상황이 생긴다. 또한, 데이터 가공에 능숙하지 않은 연구자들은 이러한 표본들을 사용할 수가 없다. 노동패널에는 중도에 패널에서 탈락하였다가 다시 진입한 표본들이 다수 있는데, 이 경우도 위와 같은 문제가 발생하여 연구자는 주의를 기울여야 한다.

### 제3절 출산력 패널자료 구축방법 검토

이 절에서는 노동패널의 다양한 출산이력 관련 변수들을 활용하여 1~19차에 포함된 표본들에 대한 출산이력 변수 구축 과정을 설명하며 이 과정에서 발생하는 어려움에 대하여 논의한다. 특히, 연구에서 흔히 사용되는 변수인 총 자녀수 변수를 패널자료로 생성하는 과정을 노동패널 홈페이지에 안내된 대로 사교육 문항을 이용하는 방법으로 구축하여 보고 발생하는 어려움과 한계점을 파악한다. 그리고 사교육 문항 외의 변수들을 사용하여 총 자녀수 변수를 생성하는 방법을 간단히 소개하고 비교하며, 여전히 남는 한계점이나 어려움에 대하여 논의하여 본다.

먼저 패널 데이터 구축을 위해 가구조사의 사교육 문항에서 총 자녀수 문항만 활용되는 것을 고려할 수 있다. 사교육 문항에서는 고등학교 이하 자녀수를 직접적으로 묻고 있어 데이터를 그대로 사용해도 되는 것처럼 보인다. 하지만, 사교육 문항의 자녀수 정보를 있는 그대로 사용하는 것은 몇 가지 문제점이 있어, 완성된 자녀수 정보에 상당한 오류가 포함되

26) 원칙적으로는 이 경우에도 추적이 되어야 하지만, 노동패널에서 혼인력 관련 변수는 응답자가 자의적으로 기분에 따라 응답을 달리하는 등 출산력변수보다 부정확할 수 있으며, 이혼가정인 경우 분가된 가구 추적이 어렵다. 또한, 분가 가구를 이어야만 자녀수를 아버지에게 매칭 시킬 수 있어 복잡하고 어려운 작업이 된다.

므로 자제하여야 하는 방식이다. 일차적으로 사교육 문항은 가구단위의 문항이므로 자녀수 문항만을 사용해서 간편하게 가구 아이디를 공유하는 특정 연령대의 집단에게 붙일 수는 있다. 그런데 이때 가구원 중 자녀의 부모가 아닌 경우가 다수 포함될 수 있어서 이 방법은 지양하여야 한다. 따라서 총 0세~고등학교(재수생) 이하 자녀수 정보(h\_1502)는 가구단위의 분석을 하는 경우를 제외하면 그대로 사용할 수 없다.

개인단위의 연구를 수행할 경우, 개인이 출산 혹은 입양한 자녀수를 추정하는 변수가 필요한데, 이때는 패널 진입 시 처음이자 마지막으로 응답한 출산경험 여부 관련 문항들(p\_9071-p\_9075)을 기본정보로 두고 사교육 문항에서 추가 출산(혹은 초산)된 자녀를 매년 파악하여 첫 응답한 자녀수에서 더해주는 방식으로 패널변수를 생성할 수 있다. 사교육변수와 개인 출산경험 변수를 모두 활용하여 매 차수마다 총 자녀수 변수를 생성하는 방법은 여러 가지가 있을 수 있지만, 본 절에서는 매 차수별로 추가로 출산되는 자녀를 찾아내는 방식을 이용해 본다. 그 이유는 추가 출산을 파악해서 더하는 방식을 사용할 경우 사교육 조사 변수상 더는 조사되지 않고 사라지는 자녀들(고등학교 졸업 후 취업, 군입대 등)로 인해 총 자녀수 변수의 값이 특정 연도에 감소하는 현상을 예방할 수 있기 때문이다. 만약 다른 방식으로 데이터를 구성하여 자녀수가 감소하는 현상이 발견되면, 이때 사라지는 자녀가 어떤 이유로 사라졌는지(예-사망, 군입대, 분가, 부모의 이혼 등)를 파악하고 총 자녀수를 재확인하는 과정이 상당히 복잡하다<sup>27)</sup>.

한편, 개인단위의 자녀수를 사교육변수를 사용하여 패널 변수로 생성할 때, 엄마의 경우만 자녀수를 생성하는 것이 바람직하다. 아빠의 경우 출산경험 관련 문항이 11차부터 조사되지 않았으므로 사교육 문항만을 이용하여서는 아빠의 자녀수를 파악하는 것이 불가능하기 때문이다. 또한, 사교육변수와 개인 출산경험 관련 변수만을 사용하여 총 자녀수 패널 변수를 생성하려고 할 때, 2차 조사에서는 사교육 문항이 조사되지 않았

27) 매년 추가 출생 자녀수를 파악하는 방식에서도 자녀가 사망하여 감소하였을 경우는 별도의 절차를 거쳐 확인하는 과정이 필요하지만, 실제 출생한 자녀의 수는 영향받지 않는다.



으므로 3차 이후 신규 진입한 표본들만 사용하는 것이 가장 편리한 방법이다. 그렇지만 표본수를 충분히 확보하기 위하여 1차년도 표본부터 사용하고자 한다면, 2차년도의 자녀수 정보가 3차년도 사교육 조사내용을 활용해 대체해야 한다. 그러므로 2차에서 탈락하는 표본들이나 2차년도에만 관측되는 표본들은 사용할 수 없다.

## 1. 총 자녀수 패널 변수 생성방법의 예

<표 4-4>는 앞서 논의한 대로 사교육 문항과 개인 출산경험 문항을 활용하여 생성할 수 있는 패널 변수인 총 자녀수 변수의 생성방법을 정리하고 있다. 기본적으로 1~19차 개인용 조사 데이터를 수직으로 연결(append)시킨다. 그다음 p\_9074(아들 수)와 p\_9075 변수를 더해주어 패널 진입 당시의 총 자녀수 변수(num\_children)를 생성한다. 다음, 3~19차 가구용 조사를 세로로 연결(append)시킨다. 그리고 자녀의 가구원번호 변수(h\_1511, h1611, h1711, h1811, h1911)를 사용하여 자녀의 생년월일과 만 나이(h\_0301~h\_0355)를 가로로 붙여준다.<sup>28)</sup> 그다음 자녀의 엄마 가구원번호 변수(h1512, h1612, h1712, h1812, h1912)를 사용하여 각 자녀별로 자녀의 엄마 가구원 아이디(mom\_pid) 변수를 생성시켜준다(그림 4-1 참조).

다음으로, 가구번호(hhid)를 이용하여 개인 패널(1~19차)과 가구용 조사(3~19차)를 가로로 연결시킨다(merge). 그 후, 개인 아이디(pid)가 가구의 엄마 가구원 아이디(mom\_pid)와 일치하는 경우가 발견될 때 child1이라는 변수값이 1을 갖도록, child1, child2, child3, child4, child5(자녀번호가 5번까지 응답 가능하므로) 생성하고, child1\_yr, child1\_mo, child1\_day와 같이 각 자녀별 생년, 생월 변수를 새로 생성하고 붙여준다. 그렇게 되면 [그림 4-2]와 같은 형식의 데이터가 생성된다.

28) 가구원의 생년월일은 10차까지 양력 음력으로 모두 응답할 수 있게 되어있는데, 음력으로 응답한 생년월일에 대해서는 일일이 양력 생년월일로 수정하여야 사용할 수 있는 문제가 발생한다. 이 과정은 처리하는 것이 매우 복잡하여 무시할 것이며, 이로 인해 데이터 생성시 오류가 발생한다.

[그림 4-1] mom\_pid 생성방법 코드 일부 발췌

```

local l=1
foreach pid of varlist h0221-h0235 {
  rename `pid' pid`l'
  local ++l
}

foreach n of numlist 1/5 {
  gen chpid`n'=.
  gen mompid`n'=.
}
foreach n of numlist 1/15 {
  replace chpid1=pid`n' if h1511==`n'
  replace mompid1=pid`n' if h1512==`n'
}
foreach n of numlist 1/15 {
  replace chpid2=pid`n' if h1611==`n'
  replace mompid2=pid`n' if h1612==`n'
}
foreach n of numlist 1/15 {
  replace chpid3=pid`n' if h1711==`n'
  replace mompid3=pid`n' if h1712==`n'
}
foreach n of numlist 1/15 {
  replace chpid4=pid`n' if h1811==`n'
  replace mompid4=pid`n' if h1812==`n'
}
foreach n of numlist 1/15 {
  replace chpid5=pid`n' if h1911==`n'
  replace mompid5=pid`n' if h1912==`n'
}

```

그 다음 개인과 차수를 기준으로 정렬한 데이터에서 <표 4-4>에서 설명하는 대로 1차부터 매 차수별로 업데이트되는 자녀수를 더해주는 방식으로 총 자녀수 변수를 생성할 수 있다. 1, 2차 진입표본을 일단 제외하고, 예를 들면, 각 표본별로 처음 패널에서 발견될 때 응답한 아들수(p\_9074)와 딸수(p\_9075)를 더해 생성된 num\_children 변수가 있을 것이다<sup>29)</sup>. 예를 들어 9차년도 첫 패널 진입자이면서 자녀가 있었던 경우를

29) 이때, num\_children값이 존재하지 않는(패널 처음 진입 당시 출산을 경험한 적이

[그림 4-2] 생성된 데이터 형태 예시

pid	hhid	year	mompid1	num_children	chage1	chpid1	child1_yr	child1_mth	child1_day	chpid2
32803	328	2000	32803	.	12	32804	1988	9	19	32805
32803	328	2001	32803	.	13	32804	1988	9	19	32805
32803	328	2002	32803	.	14	32804	1988	9	19	32805
32803	328	2003	32803	.	15	32804	1988	9	19	32805
32803	328	2004	32803	.	16	32804	1988	9	19	32805
32803	328	2005	32803	.	12	32805	1993	10	27	32804
32803	328	2006	32803	.	18	32804	1988	9	19	32805
32803	328	2007	32803	.	14	32805	1993	10	27	.
32803	328	2008	32803	.	15	32805	1993	10	27	.
32803	328	2009	32803	.	16	32805	1993	10	27	.
32803	328	2010	32803	.	17	32805	1993	10	27	.
32803	328	2011	32803	.	18	32805	1993	10	27	.
32803	328	2012	.	.	.	.	.	.	.	.
32803	328	2013	.	.	.	.	.	.	.	.
32803	328	2014	.	.	.	.	.	.	.	.

생각해 보자. 그 표본의 경우 9차년도에 `num_children` 변수의 값이 존재할 것이고 그다음 차수부터는 missing 처리되어 있을 것이다. 이러한 `num_children` 값을 개인 단위의 패널에서 매년 최대치 값을 갖도록 변수를 변경하여, 개인 단위로 보았을 때 설문 차수별로 `num_children` 값이 최소 0 (자녀출산 경험이 패널 진입 시 없었던 표본) 값을 갖도록 만들어 준다. 그 다음, 생일이 첫 번째 연도(9차) 조사년월 이후이고 두 번째 발견 차수(10차 혹은 그 후)<sup>30)</sup> 9차년(2006년)도 조사월(p\_9501) 이전인 자녀가 `child#_yr`, `child#_mo`, `child#_day` 변수를 검토하였을 때 존재할 경우 `add_children_v1`이라는 변수를 생성하고 조건에 맞는 자녀 숫자의 값을 가지도록 지정한다. 총 자녀수, `num_children` 변수의 값은 개인의 첫째(9차년도)에는 `num_children` 값을 그대로 가지고, 그다음 차수부터는 이전 차수에 발견되는 `num_children` 값에 그 차수에 계산된 `add_children_v1` 값을

없는 표본) 경우는 `num_children` 값을 0으로 대체시켜 주어야 한다.

30) 패널에 중도 탈락했다가 재진입하는 케이스도 커버된다.

더해주면 자녀수가 추가될 때마다 num\_children값이 업데이트된다.

참고로 추가 자녀수변수, add\_children\_v1을 생성할 때, 2차년도는 3차년도 사교육 문항의 정보를 활용하여 impute한 값으로 생성함은 물론 따로 처리해야 한다. 이 경우, 3차년도에 조사되지 않는 표본은 분석에서 제외된다. 3차년 add\_children\_v1변수는 1차년도 조사일자와 2차년도 조사일자 사이에 출생된 자녀수의 변수 add\_children\_2와 2차년도 조사일자와 3차년도 조사일자 사이에 출생된 자녀수 변수 add\_children\_3변수로 분리해 생성한 다음, 2차년도 add\_children\_v1값은 add\_children\_2의 값을 가지고, 3차년도 add\_children\_v1값은 add\_children\_3의 값을 가지도록 하여야 한다.

〈표 4-4〉 사교육 문항과 개인 출산경험 문항을 활용하여 출산력 데이터 구축방법 예시

차수	총 자녀수 “Num_children”		자녀수의 변화 “add-children_v1” 추적
	첫 진입자 (A)	기존 샘플 (B)	기존 샘플 (C)
	중도 탈락 후 재진입 없이 계속 관측되는 가구와 개인		
1차	개인서베이 p9071~5 출산경험 변수 활용 - 출산경험 유무에서 출산한 경험이 있다(값:1)를 응답한 개인의 아들수(p_9074)+딸수(p_9075)를 더하여 주는 방식으로 “Num_children” 변수를 생성	해당 없음	
2차	- 1차년도 (A)값 생성 방식과 동일한 방식으로 “Num_children” 변수의 값을 붙여줌	“Num_children”값에 3차년도에서 발견되는 자녀의 출생시점이 2차년도 조사시점과 1차년도 조사시점 사이라면, 해당 자녀의 수를 계산하여 붙일 수 있음. - 3차년도 (C)항에서 추정되는 “add_children_2”값을 더해줌	- 사교육 문항을 조사하지 않았으므로 사교육 문항을 활용하여 추가 자녀수를 파악할 수 없음

〈표 4-4〉의 계속

차수	총 자녀수 “Num_children”		자녀수의 변화 “add-children_v1” 추적
	첫 진입자 (A)	기존 샘플 (B)	기존 샘플 (C)
	중도 탈락 후 재진입 없이 계속 관측되는 가구와 개인		
3차	- 1, 2차년도 (A)값 생성 방식과 동일한 방식으로 “Num_children”변수의 값을 부여함	전년도의 (A)“Num_children”값에 (C)“add_children” 값을 더해 줌	개인패널에 가구아이디(hhid)를 기준으로 1:m으로 가구 데이터를 붙인다. 사교육 문항에서 0~고등학생(재학생) 이하 자녀의 pid를 찾고 각 자녀별 생년월일이 1차년도 조사일과 2차년도 조사일 사이인 자녀들에 대해 통계 패키지의 루프를 활용하여 해당 자녀별로 개인 pid가 그 자녀의 엄마 pid와 일치하는 사람들에게 자녀수만큼 add_children_01_2, add_children_02_2 등으로 변수를 생성하고, 해당 자녀가 발견될 때 1의 값을 가지도록 함. 3. “add_children_2”=add_children_02_2+add_children_02_2...의 방식으로 add_children_2 변수 생성. 3. 사교육 문항에서 0~고등학생(재학생) 이하 자녀의 pid를 찾고 각 자녀별 생년월일이 2차년도 조사일과 3차년도 조사일 사이인 자녀들에 대해 1.과 2.의 방식으로 “add_children_3” 변수를 생성 4. add_children=add_children_2+add_children_3으로 설정
4~19차	- 1, 2차년도 (A)값 생성 방식과 동일한 방식으로 “Num_children”변수의 값을 부여함	전년도의 (A)“Num_children”값에 (C)“add_children” 값을 더해 줌	1. 사교육 문항에서 발견되는 자녀들 중 생년월일이 직전 차수의 조사일과 당해 차수 조사일 사이인 자녀를 파악 2. 개인의 pid가 해당 자녀(들)의 엄마 pid와 일치하는 경우에 대하여 add_children_01, add_children_02 등으로 변수를 생성하고 조건이 일치하는 자녀가 있을 때 1의 값을 가지도록 변수들을(발견 자녀수 만큼) 생성 3. add_children=add_children_01+add_children_02+...로 계산하여 줌

## 2. 남성의 총 자녀수 변수 생성방법의 예

앞서 여성의 총 자녀수 변수 생성방법의 예에 대하여 설명하였다. 출산 변수를 가장 활발하게 사용하는 연구들은 일·가정 양립 정책 등 저출산 정책, 성별 임금격차 정책에 관한 연구들인데, 이러한 정책들의 정책 대상이 여성에만 국한되어 있다가 최근 들어 남성을 포함하는 정책으로 발전되고 있다. 예를 들어 육아휴직 정책의 경우 남성의 육아휴직과 육아휴직 사용을 증가시키는 방향으로 변화하고 있다. 따라서 남성에게도 정확한 자녀수 변수를 구축할 수 있어야 노동패널의 활용도가 높아진다.

하지만 앞서 2절에서 설명한 바와 같이, 노동패널 11차 조사 이후로는 개인조사에서 남성에게 출산경험 여부를 묻지 않고 있고, 사교육 문항에서도 자녀와 엄마를 연결짓는 문항만 있으므로 아빠의 자녀수를 정확하게 추측하는 것은 거의 불가능하다. 그렇지만 기혼남성이고 아내가 노동패널에 포함되어 있는 부부의 경우 아내의 총 자녀수 변수의 값을 남편에게 붙여줄 수 있어, 표본을 한정하였을 때 남성의 자녀 정보 변수도 활용할 수 있다.

먼저 앞선 섹션에서 예로 든 여성의 총 자녀수 변수를 생성하였다고 가정하자. 개인 패널에서 여성인 경우 개인 패널 변수인 가구주와의 관계(p\_0102)의 값(그 개인의 가구주와의 관계값)의 모든 값별로(표 5 참조) 그 값의 배우자 값과 일치하는 값을 가지는 가구원의 가구원번호를 찾기 위하여 가구조사상 가구주와의 관계변수(h\_0261~h\_0275)를 모두 검토한다. 배우자가 가져야 할 값과 일치하는 값을 가지는 가구원의 가구원번호에 해당하는 사람의 가구원 pid(h\_0221~h\_0235)를 찾고 그 값을 붙여 넣은 변수 hus\_pid를 생성한다. 예를 들어 어떤 여성 표본의 p\_0102 값이 (11)이라면 h\_0261~h\_0275변수 중 (21)의 값을 가지는 변수가 h\_0264(가구원 4의 가구주와의 관계)였다면, hus\_pid는 h\_0224 (가구원 4의 pid)와 같도록 생성한다.

다음으로, hus\_pid, year(혹은 wave), num\_children만 남긴 데이터 셋을 새로 생성하고 hus\_pid를 pid로 이름을 변경하고 num\_children 변수도 다른 이름(예-num\_children\_hus)으로 변경한 후, 기존 데이터에 pid

〈표 4-5〉 가구주와의 관계 코드

(1) 가구주 할아버지(남)	(49) 배우자의 아홉째 형제자매
(2) 가구주 할머니(여)	(51) 가구주 첫째 형제자매의 배우자
(3) 가구주 배우자의 할아버지(남)	(52) 가구주 둘째 형제자매의 배우자
(4) 가구주 배우자의 할머니(여)	(53) 가구주 셋째 형제자매의 배우자
(5) 가구주 아버지(남)	(61) 배우자의 첫째 형제자매의 배우자
(6) 가구주 어머니(여)	(62) 배우자의 둘째 형제자매의 배우자
(7) 가구주 배우자의 아버지(남)	(63) 배우자의 셋째 형제자매의 배우자
(8) 가구주 배우자의 어머니(여)	(111) 가구주 첫째자녀의 첫째자녀
(10) 가구주	(112) 가구주 첫째자녀의 둘째자녀
(11) 첫째자녀	(113) 가구주 첫째자녀의 셋째자녀
(12) 둘째자녀	(114) 가구주 첫째자녀의 넷째자녀
(13) 셋째자녀	(115) 가구주 첫째자녀의 다섯째자녀
(14) 넷째자녀	(121) 가구주 둘째자녀의 첫째자녀
(15) 다섯째자녀	(122) 가구주 둘째자녀의 둘째자녀
(16) 여섯째자녀	(123) 가구주 둘째자녀의 셋째자녀
(17) 일곱째자녀	(131) 가구주 셋째자녀의 첫째자녀
(18) 여덟째자녀	(132) 가구주 셋째자녀의 둘째자녀
(19) 아홉째자녀	(141) 가구주 넷째자녀의 첫째자녀
(20) 배우자	(142) 가구주 넷째자녀의 둘째자녀
(21) 첫째자녀의 배우자	(143) 가구주 넷째자녀의 셋째자녀
(22) 둘째자녀의 배우자	(151) 가구주 다섯째자녀의 첫째자녀
(23) 셋째자녀의 배우자	(152) 가구주 다섯째자녀의 둘째자녀
(24) 넷째자녀의 배우자	(161) 가구주 여섯째자녀의 첫째자녀
(25) 다섯째자녀의 배우자	(211) 가구주 첫째자녀의 첫째자녀의 배우자
(26) 여섯째자녀의 배우자	(212) 가구주 첫째자녀의 둘째자녀의 배우자
(31) 가구주 첫째 형제자매	(213) 가구주 첫째자녀의 셋째자녀의 배우자
(32) 가구주 둘째 형제자매	(221) 가구주 둘째자녀의 첫째자녀의 배우자
(33) 가구주 셋째 형제자매	(222) 가구주 둘째자녀의 둘째자녀의 배우자
(34) 가구주 넷째 형제자매	(223) 가구주 둘째자녀의 셋째자녀의 배우자
(35) 가구주 다섯째 형제자매	(311) 가구주 첫째 형제자매의 첫째자녀
(36) 가구주 여섯째 형제자매	(312) 가구주 둘째 형제자매의 둘째자녀
(37) 가구주 일곱째 형제자매	(313) 가구주 셋째 형제자매의 셋째자녀
(41) 배우자의 첫째 형제자매	(997) 기타 친인척
(42) 배우자의 둘째 형제자매	(998) 인척관계가 아닌 동거인
(43) 배우자의 셋째 형제자매	(1111) 가구주 첫째자녀의 첫째자녀의 첫째자녀
(44) 배우자의 넷째 형제자매	(1112) 가구주 첫째자녀의 첫째자녀의 둘째자녀
(45) 배우자의 다섯째 형제자매	

와 year를 기준으로 병합(merge)시켜준다. 그러면 남편의 pid에 num\_children 변수가 붙게 된다.

남편에게 자녀수를 붙여 준 데이터를 사용할 때는, 아내가 노동패널에서 연결된 경우만 자녀수 변수가 값을 가지기 때문에, 분석대상을 정확히 한정하고 분석을 하여야 한다. 위와 같은 경우로 남편의 자녀수를 구축하였을 시는 분석대상이 부부가 매칭이 되는 남녀(hus\_pid가 존재하는 여성과 자녀수를 연결할 때, hus\_pid가 pid로 병합이 된(\_merge값이 3인) 남성만을 기준으로 분석하여야 하고, 이러한 남편집단에서 num\_children값이 관측되지 않는 경우(missing)는 자녀가 없는 경우로 보고 연구자의 분석 목적에 따라 0의 값을 부여할 수 있다.

### 3. 기타 총 자녀수 패널 변수 생성방법 검토

앞서 총 자녀수 패널 변수를 생성하는 방법의 예시를 제시하였다. 하지만 노동패널에서 자녀수 변수를 생성하거나 검토하기 위하여 사용할 수 있는 변수가 다양하기도 하고, 사교육 문항의 자녀 정보를 사용하는 방법도 다양하여, 총 자녀수 변수를 생성하는 방법은 여러 가지이다. 이 섹션에서는 위의 예시와 다르게 총 자녀수 변수를 생성할 수 있는 방법들에 대하여 간략히 논의하여 본다.

먼저, 앞서 2절에서 간략하게 언급하였는데, 사교육 문항을 이용하되, 매년 새로 태어나는 자녀수를 추적하는 방식이 아니라 사교육 문항에서 발견되는 총 자녀수와 최초 응답한 자녀수를 비교하여 총 자녀수 변수를 생성하는 방법을 고려할 수 있다. 즉, 신규 진입 시 출산 경험이 없었던 개인에게는 총 자녀수 값 0을 부여하고, 신규 진입 시 아들수와 딸수가 관측되었던 개인은 총 자녀수를 아들수와 딸수를 더한 값으로 설정한 뒤, 신규 응답한 값을 전 패널 차수에 붙여 넣어 준 변수를 num\_children\_v3a 이라는 이름으로 생성한다. 그리고 매년 사교육 문항에서 발견되는 자녀수(자녀1~자녀5를 더한 값<sup>31)</sup>)을 가지는 변수를 생성하게 되면

31) 사교육 문항에서 0세~고등학생 미만 자녀수 변수가 있기는 하지만, 평균값이 자녀수를 각각 더한 것과 차이가 나는 경우가 있어, 여기서는 발견되는 자녀수 전



(num\_children\_v3b) 개인마다 가구단위로 사교육 문항에서 계산되는 총 자녀수 변수가 생성된다. 그다음, 매년 num\_children\_v3a값과 num\_children\_v3b값을 비교하여 다음의 규칙으로 num\_children 변수를 생성할 수 있다.

```
- num_children_v3= num_children_v3a
    if num_children_v3b<num_children_v3a
- num_children = num_children_v3b
    if num_children_v3b>=num_children_v3a
```

이렇게 num\_children변수를 생성하는 것은 사교육 문항에서 발견되는 자녀수가 첫 패널 진입 당시 응답한 자녀수보다 많을 때만 자녀수가 새로운 정보로 업데이트 되는 방식이며, 앞서 예로 든 방법보다 훨씬 간단하다.

하지만 이 방법을 사용했을 때, 출생자를 추적하는 방식이 아닌 관계로 출산이 발생하였음에도 출산의 경우를 정확하게 구분해 내지 못할 가능성이 있다. 또한 자녀의 엄마가 아닌 이모나 할머니와 같은 다른 관계에 자녀수가 연결될 수도 있다. 노동패널 데이터를 사용하는 이용자는 흔히 패널 분석 모형을 사용하기 위하여 노동패널을 사용하게 되며, 특히 개인 고정효과 모형을 사용한 분석을 많이 하게 된다. 이 때 고정효과모형에서 총 자녀수 변수는 신규 출산이 발생하였을 경우에만 변화하여 이 변화를 이용한 분석을 하게 된다. 그런데, 이 방법을 사용하여 출산이력 데이터를 구축할 경우, 신규 출생(자녀수의 변화) 정보가 앞서 예로 든 방법보다 오류가 많을 것으로 예상된다.

구체적으로, 앞서 간략히 언급하였던 자녀수가 감소하는 현상이 발생할 수 있는데, 이 점은 자녀가 사망한 경우는 드물 것이므로 자녀의 수가 증가할 경우에만 최초 자녀수 응답수를 대체하는 방식을 사용하기 때문이다. 이 같은 자녀수 감소 현상을 예방할 수는 있지만, 신규 출생 자녀를 정확하게 구분하지 못하는 문제는 여전히 남는다.

하나의 예로 한 가정에 막둥이가 태어났는데, 그 해에 첫째 자녀가 고

---

체를 더하는 값을 이용한다.

등학교를 졸업하고 취직을 하게 되거나 엄마의 재혼으로 인해 분가(아버지나 조부모 가정으로 편입됨)하거나, 다양한 이유로 사교육 조사에서 누락되면, 자녀수가 실제로는 증가하였음에도 불구하고 num\_children\_v3의 값은 그 전 해의 num\_children\_v3값과 동일하여 num\_children\_v3변숫값이 업데이트되지 않을 것이다. 연구자는 이러한 특이 케이스를 무시할 수 있지만, 문제는 특이 케이스별로 어떤 이유로 발생하는지 원인을 일일이 파악하고 각각의 케이스들을 어떻게 처리할 것인지 자의적으로 판단하여야 하는데, 변수 생성 시 사용되는 변수들이 많아 특이 케이스를 확인하는 과정 자체가 굉장히 복잡하다. 또한 연구자가 이 변수를 사용해서 개인고정효과모형을 이용한 분석을 한다면 신규 출생 시에만 총 자녀수의 변화가 생겨 ‘출생’에 의한 변화를 이용하는 분석을 하여야 하는데, 신규 출생이 발생하는 경우가 많지 않기 때문에 분석에 있어 자녀수 변화에 관한 오류가 심각한 문제를 야기할 수 있다. 마지막으로, 이 방식을 사용할 경우 3~19차 데이터만 사용할 수 있다는 한계점이 있다.

노동패널 공식 홈페이지에서 제시한 방법은 아니지만, 신규가구원 추가의 이유 변수(h\_0481~h\_0495)를 이용하여 출산이나 입양으로 인해 새로 생긴 자녀를 추적하는 방법도 있다. 방법은 앞서 예시로 든 총 자녀수 변수 생성방법처럼, 먼저 개인조사 출산이력 변수(p\_9071~p\_9075)를 활용하여 num\_children변수를 생성한 다음 add\_children 변수를 생성할 때 사교육 조사가 아닌 신규가구원 정보를 사용하는 것이다. 2차년도부터 h\_0421~h\_0435 변수(가구원 추가 탈락 여부 : 가구원 1~가구원 15) 중 2의 값(신규가구원)을 가지는 경우가 있을 때, 해당 가구원이 가구에 편입한 경로가 출산 혹은 입양인지를 h\_0481~h\_0495 변숫값이 “1-출생” 혹은 “2-입양”인지 확인한다. 만약 h\_0481~h\_0495 변숫값이 1 혹은 2일 경우, 해당 가구원의 가구주와의 관계(h\_0261~h\_0275)를 확인하여 그 관계가 엄마인 사람의 가구주와의 관계값을 모든 경우별로 구분하여 엄마로 판단되는 사람의 pid를 가진 개인에게 add\_children\_01, add\_children\_02...의 변수들을 생성하고 1의 값을 부여하여 줄 수 있다.

이 방법을 사용할 경우, 모든 차수를 동일한 방식으로 처리할 수 있다는(사교육 문항 사용 시 2차년도를 처리하는 방식이 복잡) 장점이 있으면

서, 출산을 직접적으로 추적한다는 점은 앞서 섹션 1에서 제시한 방법과 동일하다. 또한, 출산이력을 남성에게도 물어본 1~10차의 경우, 아빠에게도 엄마와 동일한 방법으로 자녀수를 계산할 수 있다는 장점이 있다. 하지만, 가구주와의 관갯값을 이용하여 자녀와 엄마를 연결하는 방법은 연구자에 따라 복잡하고 편의성이 떨어진다고 여길 수 있으며, 가구주와의 관계변수나 가구원 추가 이유 변수의 코딩에러가 일부 있는 것으로 보여, 출산변수 구축 시 오류가 발생한다. 또한, 이 방법을 사용하더라도 부부의 일시적인 별거나 이혼 등의 경우 자녀를 엄마와 정확하게 연결하지 못하는 문제는 여전히 존재한다.

#### 제4절 노동패널로 구축할 수 있는 출산력 자료의 기초통계

이 절에서는 앞서 예로 든 출산력변수 생성방법을 이용하여 생성한 총 자녀수 변수들을 비교하여 본다. <표 4-6>은 제2절에서 예로 든 사교육 문항을 이용하여 신규 출생아를 계산하기 위해 생성한 `add_children` 변수를 각 연도별로 보여준다. 2~4열은 각각 0명, 1명, 2명, 3명의 자녀가 새로 태어난 경우의 값을 기록하고 있다. 이 계산에 따르면, 매년 약 88~200명의 자녀가 표본가구의 엄마에게 태어나고 있는 것이 확인된다.

비교를 위하여 3절 마지막에 소개한 신규가구원이 출생으로 인하여 가구에 들어오게 되었을 때 가구주와의 관계 코드를 이용하여 신규출생 자녀수를 추정하는 변수도 생성해 보았다. <표 4-7>은 신규출생가구원을 추적하여 만든 `add_children_v2`변수를 <표 4-6>과 같은 방법으로 정리한 것을 보여준다. 비교해 보면, 사교육 문항을 사용하였을 때보다 신규출생으로 가구원에 들어온 자녀수를 파악하는 것이 더 많은 신규출생을 구분해 낼 수 있는 것으로 보이긴 하지만, 둘의 차이가 크지는 않다.

다음은 신규출생아수를 추적하는 방법을 사용하여 생성한 총 자녀수 변수와 제3절 마지막에 소개한 다른 방법들로 자녀수 변수를 생성하였을 때 얻을 수 있는 총 자녀수 변수의 기초통계이다. 이는 <표 4-8>에서 정

〈표 4-6〉 사교육 문항을 이용하여 생성한 신규출생자녀수 변수

add_children_v1	0명	1명	2명	3명	Total
1998	6,851	0	0	0	6,851
1999	6,074	171	1	0	6,246
2000	5,698	112	3	0	5,813
2001	5,576	155	3	0	5,734
2002	5,559	110	4	0	5,673
2003	5,808	170	2	0	5,980
2004	5,885	140	1	0	6,026
2005	5,863	147	0	0	6,010
2006	5,945	145	2	0	6,092
2007	5,960	137	4	0	6,101
2008	5,882	177	4	0	6,063
2009	7,400	129	6	0	7,535
2010	7,153	200	2	0	7,355
2011	7,076	176	7	0	7,259
2012	7,100	166	6	0	7,272
2013	7,050	161	1	0	7,212
2014	6,858	88	3	1	6,950
2015	7,146	153	4	0	7,303
2016	7,288	93	4	0	7,385
Total	122,172	2,630	57	1	124,860

리하고 있다. 먼저, num\_children 변수는 개인별로 패널 진입 시 응답한 총 아들과 딸 수를 더해준 값으로, 이 값이 존재할 때만 표본을 추출하여 통합횡단면(pooled cross-section)으로 데이터를 구축하였을 때 얻을 수 있는 값이다. num\_children\_v1은 제3절 초반에 소개한 자녀수 변수 생성 방법 예시(사교육 문항 이용)를 따랐을 때 얻을 수 있는 자녀수 변수이며, num\_children\_v2는 제3절 마지막에 소개한 신규출생가구원을 추적하는 방식으로 총 자녀수 변수를 생성한 것이다. num\_children\_v3 변수는 사교육 문항에서 나타난 총 자녀수와 처음 응답한 자녀수를 비교하여 더 큰 값으로 대체하며 가구아이디를 이용하여 가구 내의 성인여성에게 자녀수

〈표 4-7〉 신규가구원 편입이유변수를 사용하여 생성한 신규출생자녀수 변수

add_children_v3	0	1	2	3	Total
1998	6,851	0	0	0	6,851
1999	6,102	142	2	0	6,246
2000	5,697	112	4	0	5,813
2001	5,550	182	2	0	5,734
2002	5,530	140	3	0	5,673
2003	5,787	177	16	0	5,980
2004	5,840	174	12	0	6,026
2005	5,848	160	2	0	6,010
2006	5,922	164	6	0	6,092
2007	5,918	168	14	1	6,101
2008	5,848	207	8	0	6,063
2009	7,364	158	13	0	7,535
2010	7,112	235	9	0	7,356
2011	7,058	189	12	0	7,259
2012	7,103	156	11	1	7,271
2013	7,046	163	4	0	7,213
2014	6,857	85	7	2	6,951
2015	7,146	148	9	0	7,303
2016	7,261	116	6	2	7,385
Total	121,840	2,876	140	6	124,862

를 붙여주는 방식으로 생성한 자녀수 변수이다.

먼저 num\_children 변수는 자녀가 있는 사람만 응답한 변수로 평균값이 num\_children\_v1, num\_children\_v2에 비해 매우 높게 나타난다. num\_children\_v3의 경우 단순히 사교육 문항에서 나타나는 자녀수와 처음 패널 진입 시 응답한 총 자녀수의 최대값을 비교한 변수이다. 이에 따라 초기 패널 진입 당시 자녀수가 관측되거나 사교육 문항에서 자녀수가 관측되는 사람만 연결된 관계로 표본수가 num\_children\_v1, num\_children\_v2에 비해 적고 평균값이 높다. 그러므로 이 방식으로 자녀수 값을 생성하려는 연구자는 자녀가 아예 없는 사람, 자녀가 아닌데 연결된 경우(이모

〈표 4-8〉 다양한 방법으로 생성한 자녀수 관련 변수들의 기초통계

Variable	N	Mean	SD	Min	Max
num_children	7,009	2.69425	1.551613	1	18
num_children_v1	124,862	1.560443	1.765679	0	18
num_children_v2	124,862	1.564848	1.750969	0	18
num_children_v3	77,526	2.35	1.57	0	18
add_children	124,862	0.0218161	0.1492828	0	3
add_children_v2	124,862	0.0254201	0.1652419	0	3

나 나이가 많은 여자 형제 등)를 어떻게 처리할지 결정하고 사용하여야 한다.

본 연구에서는 사교육 문항을 이용하여 신규 출생자녀를 추적하는 방식과 가구원 중 신규 출생으로 인해 가구에 들어온 자녀를 추적하여 총 자녀수를 업데이트하는 방식으로 총 자녀수 변수를 생성하는 것이 그나마 오류가 가장 적은 방법일 것이라고 보았다. 그 결과 생성된 num\_children\_v1과 num\_children\_v3변수를 비교해 보면, 두 변수 간 평균값은 거의 차이가 없고, 해당 변수들을 생성하기 위해 생성된 신규출생자녀수를 추적하는 변수인 add\_children\_v1과 add\_children\_v2의 경우 평균에서 약간의 차이가 있지만 매우 적다. 단, add\_children\_v2(신규 출생 가구원 추적 방식)를 사용하였을 경우 신규 출생 자녀가 더 많이 파악되므로, 총 자녀수 변수를 활용한 패널 모형 분석을 할 때는 이 방법이 더 선호될 수 있다.

## 제5절 소 결

본 연구는 노동패널에서 출산력변수를 생성할 때 직접적인 출산 문항을 사용하지 못하고 간접적으로 변수를 생성하는 한계로 인해 발생하는 변수 생성의 어려움과 문제점에 대하여 논의하였다.

노동패널은 신규조사자에 한해서만 출산경험 여부와 총 출산한 자녀수

를 묻고 그 후로는 직접적으로 묻지 않고 있다. 그래서 자녀수나 출산관련 변수를 사용해야 할 때는 가구조사에서 사교육 문항에 나타나는 자녀 정보나 신규가구원의 편입사유(출생, 입양, 혼인 등)를 이용하여 간접적으로 출산력변수를 생성하여야 한다.

간접적으로 출산력변수를 생성하는 것은 간단해 보일 수 있지만, 실제로는 매우 복잡하고 굉장히 많은 변수들을 사용해서 추가 출생 자녀를 파악하여야 하는 관계로, 각 변수별 오류들이 집적되어 연구자가 자의적으로 각 특이 케이스에 맞춰 출생자녀수를 어떻게 판단할지 결정해야 한다. 이러한 방식으로 최종 생성된 출생자녀 변수에는 많은 오류가 포함된다. 또한 데이터를 처리하는 과정이 매우 복잡하여 접근성과 편의성도 떨어진다.

노동패널의 장점은 긴 패널기간이며, 대부분의 연구자들은 패널모형을 활용한 분석을 할 수 있어 노동패널을 선호하는 경향이 있다. 그런데 출산 관련 연구들을 패널모형으로 할 때 출산변수를 설명변수로 사용하게 되면 신규 출생의 변화에 따른 종속변수의 변화를 추정하는 것이 된다. 이때 신규 출생의 변화를 정확하게 파악하지 못하고 오류가 많이 포함되는 것은 연구자에게 부담이 될 수 있다. 특히, 신규 출생의 경우 표본에서 출생이 일어나는 경우가 매우 적어(0.02%대), 몇몇 케이스에서 발생하는 오류는 저출산이나 여성 관련 연구에서 심각한 문제가 될 수 있다.

이에 본 연구에서는 노동패널 내에서 가능한 출산력 생성방법들을 검토해 보고, 여러 가지 방법들로 생성한 출산 관련 변수들을 비교하여 보았다. 그 결과 노동패널 홈페이지에서 안내하였듯이, 출산력을 생성할 때 사교육 문항을 이용하는 방식이 가장 효과적인 방식 중 하나라고 확인하였다. 단, 본 연구에서는 사교육 문항을 사용할 때 가구주와의 관계변수와 함께 추가 출산이력을 추적하고, 신규조사자일 때 응답한 출산이력 정보를 업데이트하는 방식으로 총 자녀수 변수를 생성하는 것이 가장 정확도가 높다고 보았다. 그 이유는 그렇게 하지 않고 사교육 문항에 나타나는 총 자녀수를 그대로 활용하였을 때는 자녀가 고등학교를 졸업하거나, 부부의 이혼으로 아버지나 어머니로부터 분리되는 경우 패널 자료에서 자녀수가 감소하는 것처럼 나타나는 관측치가 꽤 많이 생기기 때문이다(방법에

따라 그 수치가 다양할 수 있는데, 적지 않음). 이때, 자녀수의 감소 원인을 파악해야 하는 어려움이 발생하고, 자녀수가 감소한 표본을 제외시킬 수는 있는데, 표본을 최대한 사용하고자 할 경우에는 아쉬울 수 있다.

또한, 본 연구에서는 출산력을 생성할 때 또 다른 좋은 방법이 사교육 문항 대신 신규가구원의 편입 사유 문항을 사용하여 신규 출생자들을 파악하는 방법임을 확인하였다. 특히 이 방법은 사교육 문항을 사용하였을 때보다 더 많은 신규출생자를 파악할 수 있게 한다.

하지만 두 방법 모두 처리과정이 매우 복잡하여 데이터사용에 익숙치 못한 연구자는 거의 사용할 수 없는 방법으로 판단되며, 가구주와의 관계 등 다양한 변수를 사용하여야 하는데 각 변수들이 오류를 포함하고 있는 것으로 보여, 여전히 정확도가 떨어지는 것으로 판단된다.

한편, 노동패널 출산력 데이터의 또 다른 한계점은, 남성의 자녀 수가 정확도 면에서 매우 떨어진다는 점이다. 남성의 경우 10차 조사 이후 신규조사자에게 출산경험을 묻지 않아, 신규조사자 출산 경험 변수는 사용할 수 없고 오로지 가구주와의 관계를 이용하여 자녀수를 파악하여야 한다. 따라서 여성의 경우보다 출산력 정보의 정확도와 편의성이 더 떨어진다.

노동패널 사용자의 편의성을 향상시키고 출생력 데이터의 정확도를 높이기 위해서는 다음과 같은 개선 방안을 생각할 수 있다. 우선, 다음 차수 조사 때부터 혼인이력의 변화를 직접 질문하듯 매년 출산이력의 변화를 직접 질문하도록 하는 것이 필요해 보인다. 또한, 이전 차수의 출산이력을 구축하는 과정들은 매우 복잡하므로, 이미 조사된 차수들에 대해서는 간접적으로 조사된 모든 변수들을 활용하여 매년 예상되는 총 자녀수를 개인 pid별로 기록한 ‘출산력’ 데이터를 구축하여 사용자가 필요에 따라 사용할 수 있도록 할 수 있다. 이런 방식으로 노동패널 출산력 데이터를 개선할 경우, 사용자의 접근성과 편의성이 향상될 것으로 보이고 따라서 관련 연구가 활발히 진행될 수 있을 것으로 보인다.



## 제 5 장

### 사업체패널조사의 경영의사결정 참여 설문문항들의 효과적 활용 탐색

#### 제1절 서 론

본 연구는 2005년부터 2015년까지 6차례에 걸쳐 조사된 사업체패널조사의 노동조합 또는 노사협의회 경영의사결정 참여 설문문항들의 활용 방법 탐색을 목적으로 한다. 사업체패널조사는 노동조합 현황, 단체교섭, 노동쟁의, 경영 참여, 노사협의회 및 고충처리, 그리고 작업장 혁신 영역에 관한 설문조사를 회사의 노사관계 담당자와 노동조합 대표자를 대상으로 진행하고 있다. 이 영역들을 측정하기 위한 설문문항들은 매 조사마다 변화해왔지만 기본적으로 일관된 틀을 유지하고 있다. 이는 경영의사결정 참여에 관한 설문문항에도 해당된다.

본 연구는 근로자 대표조직의 경영 참여를 측정하는 총 11개 설문 문항들의 신뢰도와 타당성을 검토하고자 한다. 경영 참여 설문문항들이 측정하고자 하는 근로자 대표조직의 경영 참여가 가지는 이론적·현실적 중요성에도 불구하고 사업체패널조사 내 노사관계 영역들의 다른 설문문항들에 비해 그 활용도가 높지 않다. 저자가 알기에는 지금까지 이 설문문항들을 사용하여 분석한 논문이 2편에 불과하다. 그 이유로는 아마도 1) 사용된 척도에서 “97) 해당없음”의 모호함, 2) 등간척도로 사용하기 어

려움, 3) 연도별 요인구조의 불안정성을 들 수 있을 것이다(박경원, 2016).

본 연구는 이러한 이유 (1)과 (2)를 “해당없음”을 미참여로 간주하고 사용된 척도들을 기반으로 근로자 대표조직의 경영 참여를 1) “노사협약”과 “공동결정”, 그리고 2) “공동결정”과 같이 광의와 협의의 경영 참여로 정의함으로써 직접적으로 다루고자 한다. 이렇게 정의된 경영 참여를 기반으로 신뢰도 및 타당성을 검증한다. 특히 사업체의 경영 참여 수준을 측정하는 참여지수를 계산하고, 경영 참여 수준의 결정변수들과 성과변수들을 패널고정효과모형으로 타당성을 검증한다. 총 11개 영역에서의 참여에 동일한 가중치를 부여하는 단순합, 그리고 영역별 참여 난이도에 따른 가중치를 부여하는 Rasch 추정값을 동원해 참여지수를 구성하였다.

본 연구의 실증분석은 크게 세 부분으로 구분할 수 있다. 첫째, 경영의 사결정 참여 설문문항들에 대한 신뢰도를 검증한다. 노사 양측을 동시에 조사한 2013년 설문문항들의 일치도와 Spearman's rho를 추정하여 평가자 간 신뢰도를 추정한다. 그리고 2005년부터 2015년까지의 설문문항들에 대한 연도별 내적 일관도를 추정한다.

둘째, 경영참여지수의 타당도를 검증하기 위하여 경영 참여 수준, 선행변수들, 성과변수들 간 상관관계를 분석한다. 선행변수들로 노사 양측에 의해 측정된 노사의 노사관계정책(예, 회사의 파트너십과 노동조합의 협조전략), 노사관계 결과변수로는 노사관계분위기와 고용안정성을, 성과변수로는 정성적·정량적으로 측정된 노동생산성과 재무성과를 포함하였다.

마지막으로, 패널고정효과모형을 이용하여 경영 참여 수준과 선행변수들, 그리고 성과변수들과의 관계를 분석한다. 이를 위해 주요 변수 및 통제변수에 결측치가 없이 3회 연속 관찰된 사업체들을 분석 대상으로 선정하였다. 구체적으로 경영 참여 수준의 결정요인으로 전체 표본을 이용하여 노동조합의 조직 여부를, 유노조사업체 표본을 이용하여 노사 양측이 응답한 회사의 파트너십과 노동조합의 협조전략을 분석하였다. 경영 참여 수준의 사업체 성과에 대한 영향을 파악하기 위하여 성과변수로서 노사 양측이 응답한 노사관계분위기와 고용안정성을, 그리고 정성적, 정량적 성과지표로 구성된 노동생산성과 재무성과를 종속변수로 분석하였다.

결론에서는 본 연구의 분석결과들과 함의를 논의한 후 향후 경영의사결정 참여 설문문항을 분석에 사용하기 위한 제안 사항들을 논의한다.

## 제2절 경영의사결정 참여 설문문항의 분석 개요

### 1. 신뢰도와 타당도<sup>32)</sup>

#### 가. 신뢰도

측정과 관련하여 가장 중요한 지표 중의 하나로서 신뢰도(reliability)는 측정하거나 관찰한 값들이 동일하거나 일관된 결과를 보이는 정도를 나타내는 개념이다. 신뢰도가 높은 설문문항은 측정값이 안정되고(stable), 일관되며(consistent), 따라서 믿을 만하며(dependable), 예측가능성(predictability)이 높다는 것을 의미한다(Kerlinger, 1964).

사업체패널조사의 경영의사결정 설문문항들의 신뢰도를 추정하기 위하여 고려할 사항은 다음과 같다. 첫째, 이 설문문항들이 사용하는 척도는 “1) 사전통보도 없이 회사가 결정, 2) 사전통보, 3) 노사협의, 4) 노사공동결정, 97) 해당없음”이라는 점이다. 연구자 고유의 척도 관련 가정에 따라 이 설문문항들을 명목, 서열, 또는 등간척도로 분석에 사용할 수 있다. 본 연구에서는 명목척도로서 다루고자 한다.

2005년부터 2015년까지 사측에 조사되었지만, 2013년에 한해 노측에게도 조사되었다. 따라서 2013년 설문문항들을 분석하여 노사의 사업체 내 참여정도 평가가 얼마나 일치하는지, 즉 평가자 간 일치도를 측정할 수 있다. 또한 척도 관련 가정을 완화하여 설문문항들을 서열척도로서 간주하고 Spearman's rho를 통해 노사 양측 응답의 일관성을 평가한다.

다음 고려사항은 설문문항수로, 경영의사결정 참여 설문문항은 총 14

32) 이 섹션은 박경원(2016)의 제2절을 경영의사결정 설문문항에 맞게 재정리한 것임

개, 전 기간 모두 측정된 설문문항수는 11개로 복합 설문문항으로 구성되어 있다. 총 11문항으로 근로자 대표조직의 경영 참여를 측정하지만, 이 구성개념이 단일 요인인지, 아니면 복수의 하위 요인으로 구성되어 있는지 명확하지 않다. 본 연구에서는 사업체의 경영 참여라는 잠재 특성을 반영하는 단일 요인으로 가정하고, 참여지수를 구성하기 위한 일관성을 가지는지 내적 일관도를 통해 평가한다.

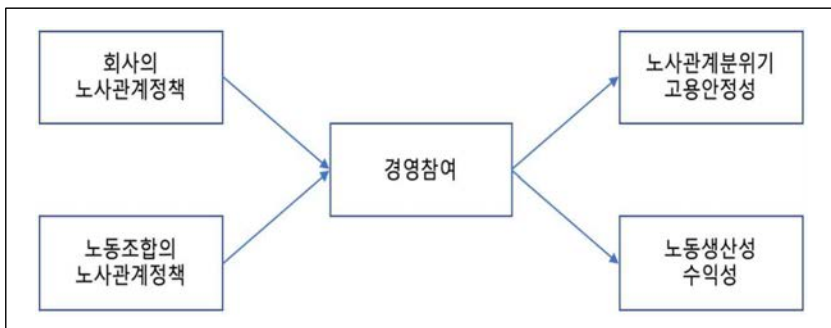
다시 강조하지만, 사회과학 분과에 따라 같은 척도일지라도 다른 방식으로 척도를 활용하기 때문에, 본 연구는 여러 가능한 방식들 중의 하나로 경영 참여를 측정하는 시도로 볼 필요가 있다.

#### 나. 타당도

타당도는 측정도구가 측정하고자 하는 대상을 충실히 측정한 정도로 정의할 수 있다. 본 연구는 경영참여변수가 이론적으로 관계가 있는 변수와의 상관관계를 측정하는 수렴(convergent) 타당도에 초점을 두며, 이 경우 상관관계가 높을수록 타당도가 높다고 할 수 있다. 한편 제한적이지만 “해당없음”을 포함하여 이론적으로 관계가 없는 변수와의 상관관계를 측정하는 판별(discriminant) 타당도를 검증한다. 이 경우 상관관계가 낮을수록 타당도가 높다고 할 수 있다.

본 연구는 근로자 대표조직의 경영 참여와 이론적으로 관계가 있는 선행변수들과 성과변수들과의 상관관계를 평가하고, 패널고정효과모형을

[그림 5-1] 타당도 검증을 위한 실증모형



통해 다른 변수들의 영향을 통제한 후에도 경영 참여 수준과 선행 및 성과변수들 간 관계가 일반화될 수 있는지를 검토한다.

[그림 5-1]은 구성개념 타당도를 검증하기 위한 실증모형을 제시한다. 이 모형의 기본 아이디어는 김정우(2018)의 분석모형에서 비롯되었지만, 다양한 변수들을 포함시키는 대신 모델을 단순화하여 실증분석틀로 활용한다.

#### 다. 동일방법편의

근로자 대표조직의 경영의사결정 참여 설문문항들의 타당도를 검증하는 데 있어서 중요한 고려사항들 중 하나는 측정변수들 간의 관계에 동일방법편의가 미치는 영향일 것이다. 동일방법편의는 응답자 또는 측정방법에 의해 발생할 수 있다(박원우 외, 2007). 동일방법편의에 의한 직접적인 결과는 변수들 간 상관관계가 실제보다 높거나 낮게 추정되는 것이며, 그 결과 Type I 에러 또는 Type II 에러의 가능성을 높이는 것이다(Siemsen, Roth, & Oliveira, 2010). 따라서 가능하면 독립변수와 종속변수를 서로 다른 응답자를 통해 측정하거나, 측정방법들을 다르게 활용함으로써 동일방법편의를 최소화할 것이 권장된다.

동일방법편의와 관련하여 사업체패널조사의 장점은 다음과 같다. 첫째, 특정 주제에 대해서는 사측의 인사·노무담당자와 노측의 노동조합 대표자가 응답하게 함으로써 상관관계 분석에서 동일방법편의의 상대적 크기를 추정할 수 있는 조건을 제공한다. 본 연구에서는 회사와 노동조합의 노사관계정책, 노사관계분위기, 그리고 2013년에 한하지만 경영의사결정 참여가 이에 해당된다. 또한 노사 양측에 동일하게 대면설문조사를 수행함으로써 측정상황을 동일하게 통제된 것으로 볼 수 있다.

둘째, 분석에 사용된 변수들을 측정하기 위한 설문문항들이 각기 다른 척도들을 사용하고 있다. 모두 응답자의 주관적 “평가 또는 판단”을 근거로 하지만, 경영의사결정 참여, 노사 양측의 노사관계정책, 고용안정성은 변수의 고유성을 반영하는 비정형적 척도들을 사용함으로써, 노사관계분위기, 노동생산성, 재무성과를 측정하는 설문문항들의 리커트 타입 척도와의 유사성을 최소화하였다.

마지막으로, 응답자의 주관적 “평가 또는 판단”으로 측정된 정성적 성과지표(예, 노동생산성과 재무성과)는 외부 DB를 통해 확보된 사업체의 정량적 성과지표(예, 1인당 부가가치와 1인당 영업이익)와 함께 제공하기 때문에 동일방법편의가 성과변수들에 미치는 영향을 비교할 수 있다.

## 2. 경영의사결정 참여 설문문항을 분석한 국내 연구

사업체패널조사의 근로자 참여 관련 설문문항들을 분석한 실증연구들은 소집단활동, QC 활동, 팀제, 제안활동과 같은 작업장 수준의 참여적 작업관행 또는 고성과작업시스템이 사업체 성과에 미치는 영향들에 집중되어 왔다<sup>33)</sup>. 반면 본 연구와 관련하여 경영의사결정 참여 설문문항들을 활용한 논문은 지금까지 두 편이 출간되었다.<sup>34)</sup>

김현동·이동진(2010)은 2007년 사업체패널조사를 분석하여 근로자 대표기구의 경영활동 참여가 작업현장 자율권과 참여적 작업조직 프로그램에 통계적으로 유의한 정(+)의 영향을 미치는 것으로 보고하였다. 이들은 근로자 대표기구의 경영활동 참여를 “경영전략”, “투자전략”, “신기술·신기계 도입”, “외주하청” 그리고 “국내외 공장 이전”과 같이 전략적 수준의 경영의사결정 영역에 대한 설문문항을 이용하여 측정하였다. 이들은 이 문항들을 등간척도로 다루면서 “사전에 통보도 없이 회사가 결정”을 1점, “사전통보”를 2점, “노사협의”를 3점, “노사공동결정”을 4점으로 코딩하였으며, 5 문항들의 평균값을 분석에 사용하였다. 따라서 이 설문문항들의 “해당없음”은 결측치로 분석에서 제외되었으나, 그 수는 보고되지 않았다.

김정우(2018)는 2005년부터 2015년까지의 사업체패널조사를 이용하여 Freeman과 Lazear(1995)의 근로자 경영참여모델을 검증하였다. 그는 이론적 모델이 예측한 대로 생산성 지표로서의 1인당 부가가치와 1인당 인

33) 사업체패널조사를 활용한 근로자 경영 참여 연구들에 대한 포괄적 리뷰는 송민수·김정우(2016)를 참조할 것.

34) 지형주·양동훈(2017)은 학술대회에서 발표되었으나 인용하지 말 것을 주문하여 리뷰에서 제외함.

건비가 경영 참여 수준이 높아질수록 일정 수준까지 상승하다가 그 이후 하락하는 역U자형의 패턴을 검증하였다.

총 11개의 경영의사결정 참여 설문문항들이 분석에 이용되었으며, 이 문항들은 김현동·이동진(2010)과 같은 방식으로 1~4점으로 코딩된 후 사업체가 응답한 문항들의 총합을 응답한 문항수로 나누어 계산된 사업체당 평균점을 참여지수로 분석하였다. 따라서 사업체가 “해당없음”으로 응답한 문항들은 사업체의 지수 계산에서 제외되었지만, 이로 인한 표본 수 감소는 없었다. 또한 패널고정효과모형에 참여지수의 지연(lagged) 변수를 사용함으로써 최소 2회 연속 관찰된 표본들을 분석에 활용하였다.

두 편의 논문들은 참여 설문문항들을 1~4점 등간척도로 간주하였으며, 각 영역의 참여에 동일한 가중치를 부여했다는 공통점이 있다. 따라서 두 편 모두 설문문항들을 1~4점 등간척도로 사용하기 위해 강한 가정을 도입하였으며, 이에 경영참여 점수를 해석하는 데 모호함이 존재한다. 예를 들어, 4개의 의사결정 영역에서 1점, 1점, 4점, 4점일 경우, 경영참여 점수가 2.5점이 되는데, 이를 “사전통보”와 “노사협의”의 중간 정도로 해석할 수 있다. 즉 분석의 단순함을 높이는 동시에 해석의 모호함도 높인 것으로, 연구자의 선호에 따라 수용하는 것은 다를 수 있다.

## 제3절 데이터 및 변수의 측정

### 1. 데이터

본 연구는 2005~2015 사업체패널조사에서 사업체 내 근로자의 대표조직, 즉 노동조합 또는 노사협의회가 조직되어 있으며, 3회 이상 연속으로 응답한 사업체들을 선정하였다. 먼저 노동조합 또는 노사협의회의 조직을 조건으로 한 이유는 근로자의 경영의사결정 참여가 사업체에서 영향을 미치기 위해서는 근로자 대표조직이 형식적이건 실질적이건 존재해야 하기 때문이다. 따라서 2개 이상의 복수노조 사업체나 제1노조의 지위 변

화 등 근로자 대표조직 내 리더십 및 운영 관련 이슈들은 경영의사결정 참여수준에 영향을 미치는 요인이기 때문에 표본 선정에는 고려하지 않았다. 다음으로 3회 이상 연속으로 조사된 사업체를 선정한 이유는 경영의사결정 참여수준의 시간적 변홑값을 적어도 2회 이상 반영되게 함으로써 사업체 내 참여수준과 성과변수들 간의 관계를 보다 안정적으로 추정하기 위한 것이다. 이러한 표본 선정과정을 통해서 노동조합 또는 노사협의회가 결측치이거나 조직되지 않은 사업체 1,375개가 제외되었으며,<sup>35)</sup> 1~2회 조사된 사업체 3,786개가 제외되었다. 따라서 총 12,353개 사업체 중에서 경영의사결정 참여수준을 기준으로 7,192개 - 2005년 1,041개, 2007년 1,219개, 2009년 1,300개, 2011년 1,412개, 2013년 1,270개, 2015년 950개 - 의 사업체들이 기본적인 분석 대상으로 선정되었으며, 통제변수와 종속변수들의 결측치에 따라 다른 수의 표본이 분석에 이용되었다.

## 2. 변수의 측정

### 가. 경영의사결정 참여 여부

사업장 내 근로자 대표조직으로서 노동조합 또는 노사협의회의 경영의사결정에 대한 참여 정도를 측정하기 위한 설문문항 번호와 내용은 <표 5-1>에 정리되어 있다. 2005년부터 2015년까지 조사된 근로자 대표조직의 경영의사결정 참여 영역은 총 14개로, 경영 및 투자 전략부터 임금제도, 노동시간 변경 등 사업체의 전략적, 기능적 활동을 포괄하고 있다. “근로시간제도 변경”, “정규직(무기계약직) 전환”, 그리고 “복리후생제도 변경”은 조사 당시 노동현안을 반영하기 위해 추가된 영역들이며, 나머지 11개 영역들은 2005년부터 일관되게 조사되고 있다. 이에 본 연구는 이 11개 영역들의 경영의사결정 참여수준을 측정한다. 사용된 척도는 “(1) 사

35) 의무적인 노사협의회 설치 기준이 상시 30인 이상의 사업체이기 때문에 30인 이상 사업체가 조사대상인 사업체패널조사에는 모든 사업체에 노사협의회가 설치되어 있어야 한다. 그러나 유노조사업체 261개, 무노조사업체 1,172개로 총 1,433개의 사업체에는 노사협의회가 설치되지 않았다. 이 중 본 연구에서 유노조사업체는 표본선정과정에 포함하였다.



〈표 5-1〉 노동조합/노사협의회의 경영의사결정 참여 설문문항

경영의사결정 영역	설문문항 번호			비고
	유노조 (사측)	무노조 (사측)	유노조 (노측)	
경영전략	PQ5002	RQ3002	PQ3915	2015년 추가
투자전략	PQ5003	RQ3003	PQ3916	
신기술·신기계의 도입	PQ5004	RQ3004	PQ3917	
외주하청	PQ5005	RQ3005	PQ3918	
국내외 공장 이전	PQ5006	RQ3006	PQ3919	
근로자의 배치 전환	PQ5007	RQ3007	PQ3920	
비정규직 채용	PQ5008	RQ3008	PQ3921	
정규직(무기계약직) 전환	PQ5009	RQ3009		
명예퇴직	PQ5010	RQ3010	PQ3922	
인사평가제도 변경	PQ5011	RQ3011	PQ3923	
임금제도 변경	PQ5012	RQ3012	PQ3924	2015년 추가
복리후생제도 변경	PQ5013	RQ3013		
승진제도 변경	PQ5014	RQ3014	PQ3925	2013년 추가
근로시간제도 변경	PQ5015	RQ3015	PQ3926	

주: 설문문항 번호는 2015년 사업체패널조사 배포용 기준임.

전통보도 없이 회사가 결정, (2) 사전통보, (3) 노사협의, (4) 노사공동결정, (97) 해당없음”이다.

근로자 대표조직의 경영의사결정 참여 수준을 측정하기 위하여 연구자가 결정할 사항은 다음과 같다. 첫째, 이 설문문항을 어떻게 조작할 것인가이다. 각 문항에 사용된 척도는 명목척도, 서열척도, 또는 강한 가정을 통해 등간척도로도 조작할 수 있다. 특히 “해당없음”을 어떻게 해석할 것인가가 중요하다. 예를 들어, 명목척도로 다룬다면 “해당없음”을 포함할 수 있지만, 서열 및 등간척도로 다룬다면 결측치로 처리해야 한다. 또한 명목 및 서열척도로 다룬다면, 각 4개와 3개의 더미변수를 생성함으로써 척도에 내재한 특성을 측정할 수 있다. 반면 등간척도로 다룬다면, 1개의 변수로 측정할 수 있다. 한편 서열 및 등간척도로 다룰 경우, “해당없음”의 빈도만큼 결측치가 발생하게 된다. 이 경우 연도별로 적지 않은 사업

체들이 표본에서 탈락하며, 패널 분석을 수행한다면 훨씬 많은 수의 표본이 탈락하게 된다. 따라서 본 연구에서는 서열 및 등간척도가 아닌 명목척도로 조작한다.

다음으로, 주어진 척도를 그대로 사용할지 아니면 단순화할지를 결정해야 한다. 전자의 경우, 예를 들어, “사전통보도 없이 회사가 결정”을 기준으로 각 카테고리들을 반영하는 4개의 더미변수를 사용하게 된다. 후자의 경우, 주어진 척도를 기반으로 참여를 1) “노사협의”와 “노사공동결정”으로, 또는 2) “노사공동결정”만으로 정의할 수 있다. 따라서 각 정의에 따라 참여를 반영하는 1개의 더미변수를 사용하게 된다. 본 연구에서는 이와 같이 참여의 강도를 구분하기 위하여 두 가지 방식으로 참여를 정의한다.

한편 여전히 “해당없음”에 대한 해석의 문제가 남는다. 아마도 응답자들은 주로 사업체 내에서 해당 경영의사결정이 발생하지 않는 경우에 “해당없음”을 선택했을 것으로 생각된다. 본사 또는 다른 사업체의 경영진이 경영·투자 전략 권한을 가지고 해당 사업체의 경영진은 결정사항을 실행만 하는 경우, 마찬가지로 본사 경영진이 모든 직원들의 임금 및 노동시간제도를 결정하는 경우를 예로 들 수 있다. 해당 사업체 외부에서 진행되는 의사결정과정에 근로자 대표조직이 참여했을 가능성도 있으나, 사업체패널조사에서는 관찰되지 않기 때문에 확인할 수 없다. 본 연구에서는 이러한 가능성이 존재하지만 해당 사업장에서 경영 참여가 발생한 것이 아니기 때문에 미참여로 간주한다. 이러한 잠재적인 측정오차가 보수적인 통계적 추론을 하도록 이끄는 것도 그 한 이유라 할 수 있다.

셋째, 위와 같이 정의된 경영 참여로 코딩된 총 11개의 경영의사결정 영역을 각각의 변수로 분석에 사용할지, 아니면 근로자 대표조직의 참여라는 사업체의 잠재(latent) 특성을 측정하는 지수(index)로 분석에 사용할지를 결정해야 한다. 전자의 경우, 총11개 영역의 참여를 측정할 더미변수들의 계수에 대한 통계적 추론이 분석의 목적이 되며, 통계적 추론 또한 매우 직접적이라 할 수 있다. 그러나 영역들 간 상관관계로 인해 계수에 대한 통계적 추론이 제약될 수 있으며, 특히 여러 영역들에서의 참여가 가질 수 있는 누적적 영향, 즉 참여수준을 통해 측정될 수 있는 시너지효과는

통계적 추론에서 제외된다. 이러한 이유로 본 연구는 참여수준을 반영하는 사업체의 잠재 특성을 측정하는 참여지수를 분석에 사용한다.

본 연구는 참여수준을 측정하는 가장 단순하고 직접적인 단순합을 분석에 사용한다. 11개 의사결정 영역의 참여에 동일한 가중치를 부여하고 계산하는 단순합 참여지수는 그만큼 직관적이며, 참여수준 현황 파악과 통계적 추론이 쉬운 장점이 있다. 그러나, 모든 의사결정 영역에 동일한 가중치를 부여한다는 점에서 적절한가라는 문제제기가 있을 수 있다. 예를 들어, 경영전략 영역과 임금제도 변경 영역의 공동결정이 실질적으로 동일한 참여인가라는 문제가 있다. 따라서, 본 연구는 단순합 참여지수뿐만 아니라 영역별 참여의 난이도를 반영하여 각 영역의 참여에 상이한 가중치를 부여하는 Rasch 참여지수를 사용한다.<sup>36)</sup>

요약하면, 본 연구는 총 11개의 경영의사결정 설문문항들을 활용하기 위하여 참여를 두 가지 방식, “노사협의”와 “노사공동결정”(코딩방식 1) 그리고 “노사공동결정”(코딩방식 2)으로 정의한다. 이러한 정의에 따라 “해당없음”은 “사전통보도 없이 회사가 결정”과 “사전통보”와 함께 미참

36) Rasch 측정모델을 사용하는 이유는 노동조합 또는 노사협의회 참여가 경영의사결정 영역별로 상대적으로 어렵거나 쉬울 수 있으며, 이러한 난이도를 반영한 가중치를 추정하여 참여수준을 추정할 수 있기 때문이다. 예를 들어, 경영의사결정 영역들에서 참여 빈도수가 상대적으로 낮은 영역의 참여에 더 높은 가중치를 부여함으로써 각 영역들의 난이도를 반영할 수 있다. 보다 기술적으로 말하자면, Rasch 측정모델은 사업체  $i$ 의 근로자 대표조직이 특정 영역에서 참여할 확률이 사업체의 경영참여 정도를 반영하는 잠재적 특성  $\theta$ 와 경영의사결정 영역  $j$ 의 난이도  $\gamma$ 에 의해 결정되는 것으로 모델링한다. 이는 다음 수식 (5-1)과 같이 정리할 수 있다.

$$P(X_{ij} = 1) = \Phi(\theta_i, \gamma_j) \quad (5-1)$$

$\theta$ 는 사업체  $i$ 의 경영참여 정도를 가리키는 잠재적 특성,  $\gamma$ 는 경영의사결정 영역  $j$ 의 난이도를 가리키는 잠재적 특성을 가리킨다.

함수  $\Phi$ 는 아래 수식 (5-2)와 같이 로지스틱 형식을 가진 것으로 정의된다.

$$P(X_{ij} = 1) = \frac{\exp(\theta_i - \gamma_j)}{1 + \exp(\theta_i - \gamma_j)} \quad (5-2)$$

사업체의 잠재적 특성  $\theta$ 와 영역의 난이도  $\gamma$ 는 Maximum Likelihood Estimation 방법으로 추정하게 된다. 본 논문에서는 STATA의 Hardouin(2007)의 rasch.ado를 이용하여 Rasch 추정값을 계산하였다.

여로 코딩되며, 이렇게 코딩된 11개의 더미변수들은 참여수준을 측정하는 참여지수, 즉 단순합과 Rasch 추정값 분석에 사용된다.

#### 나. 경영의사결정 참여수준의 선행 및 성과 변수의 측정

##### 1) 경영의사결정 참여수준의 선행변수

경영의사결정 참여와 관련된 노사관계의 선행변수로 사측과 노측 각각의 노사관계정책을 측정하였다. 사업체의 노사관계정책은 “귀 사업장에서 사측의 노동조합에 대한 정책은 어떠합니까?” 문항을 이용하여 측정하였다. 척도는 “(1) 노동조합을 약화시키거나 해체하려고 한다”, “(2) 노동조합의 실체를 인정하지만, 경영에 대한 참여는 최소화시키려 한다”, “(3) 노동조합을 파트너로 존중하고 경영에 적극 참여시키려 한다”이며, (1)과 (2)를 회피전략(0)으로, (3)을 파트너십(1)으로 구분하여 코딩하였다.<sup>37)</sup>

노동조합의 노사관계정책은 “귀 사업장에서 현 노조집행부의 노사관계정책은 전반적으로 어떻습니까?” 문항을 이용하여 측정하였다. 척도는 “(1) 사측의 경영권을 최소화하려 노력하고 전투적이다. (2) 사측의 경영권을 인정하지만, 조금이라도 노동조합과 근로자의 이익에 반할 우려가 있는 사항에 대해서는 전혀 협조적이지 않다. (3) 사측의 경영에 적극 협조함으로써 장기적으로 근로자의 이익을 지키려 한다. (4) 근로자의 이익을 크게 고려하지 않고 사측에 협조적이다.”이며, (1)과 (2)를 분리전략으로(0), (3)과 (4)를 협조전략(1)으로 구분하여 코딩하였다.

##### 2) 경영의사결정 참여수준의 성과변수

경영의사결정 참여수준과 관련된 성과변수들을 크게 정성적 지표와 정량적 지표로 구성하였다. 먼저 정성적 지표로서 노사관계분위기, 고용안

37) 사측(노측) 응답에서 노동조합 압박(suppression) 전략을 채택한 사업체 수가 2005년 3(21)개, 2007년 4(25)개, 2009년 5(13)개, 2011년 11(20)개, 2013년 3(17)개, 2015년 8(10)개로 빈도가 매우 낮았다. 이에 더 포괄적인 회피전략으로 개념화하여 측정하였다.

정성, 노동생산성과 재무적 성과를 측정하였다. 노사관계분위기는 세 개의 변수들로 측정하였다. 우선 단일문항인 “귀 사업장의 전반적인 노사관계는 어떠합니까?”이며, 사용된 척도는 “(1) 매우 나쁘다, (2) 나쁜 편이다, (3) 보통이다, (4) 좋은 편이다, (5) 매우 좋다”이다. 다음으로 노사 양측의 응답을 기반으로 한 복합 설문문항들로 측정한 노사관계분위기이다. 노사 양측에 동시에 조사된 5개의 문항들, 즉 “노사는 서로 약속을 잘 지킨다”, “협상은 노사가 상호 신뢰하는 분위기에서 이루어진다”, “노사 간 정보교환이 잘 이루어진다”, “근로조건의 중요한 변화는 대부분 노사가 공동으로 협의한다”를 이용하여 노사관계분위기를 측정하였다. 사용된 척도는 “(1) 전혀 그렇지 않다, (2) 그렇지 않다, (3) 보통이다, (4) 그렇다, (5) 매우 그렇다”이다.

고용안정성은 두 개의 설문문항들, 즉 “정규직근로자에 대한 인원감축이나 경영상의 이유로 인한 해고 여부”와 “정규직근로자에 대해 경영상의 이유로 인한 해고를 하지 않는다는 경영정책이나 합의사항 여부”를 이용하여 해고가 없으면서 합의사항이 있는 경우를 고용안정성이 있는 것으로 측정하였다.

전년 기준 동종업체 평균과 비교한 노동생산성과 재무적 성과를 사용하였다. 이 두 문항들에 사용된 척도는 “(1) 동종업체 평균에 비하여 매우 낮다, (2) 동종업체 평균에 비하여 낮은 편이다, (3) 동종업체와 비슷하다, (4) 동종업체 평균에 비하여 높은 편이다, (5) 동종업체 평균에 비하여 매우 높다, (97) 해당없음, (99) 모름”이며, (97)과 (99)는 결측치로 처리하였다.

정량적 지표로는 1인당 부가가치와 1인당 영업이익을 측정하였다. 생산성을 측정하는 변수들 중의 하나인 1인당 부가가치는 당해 연도의 부가가치를 회계기간 동안의 평균 근로자수로 나눈 값의 로그값을 사용하였다. 부가가치는 영업이익+영업외이익+인건비(급여총액+퇴직급여+복리후생비)+순금융비용+임차료+세금과공과+감가상각비로 계산하였다. 수익성을 측정하는 변수들 중의 하나인 1인당 영업이익은 당해 연도의 영업이익을 회계기간 동안의 평균 근로자수로 나눈 값의 로그값을 사용하였다. 영업이익은 당기매출액에서 영업비용을 뺀 값으로 사용하였다.

## 제4절 경영의사결정 참여수준의 계량적 분석

### 1. 경영의사결정 참여 설문문항의 신뢰도 분석

경영의사결정 참여를 측정하는 설문 항목들은 2013년에만 노사 양측에 조사되었다. 이는 사업체 내 노동조합 대표조직의 경영 참여를 바라보는 노사의 관점이 어느 정도 일치하는지, 또는 일관되는지, 즉 평가자 간 신뢰도(reliability)를 추정할 수 있는 좋은 기회를 제공한다.

설문문항들의 신뢰도를 측정하는 방법은 매우 다양하지만, 본 연구에서는 일치도(agreement)와 Spearman's rho를 이용하여 경영의사결정 참여 관련 설문의 응답에 대한 노사 간 신뢰도를 측정하였다. 먼저 경영의사결정 설문문항에 대한 노사 양측의 응답이 절대적으로 일치하는 정도를 일치도로 측정할 수 있다. 다음으로 “(97) 해당없음”을 결측치로 처리하고 “(1) 사전통보도 없이 회사가 결정, (2) 사전통보, (3) 노사협의, (4) 노사공동결정”을 참여 수준이 반영된 서열척도로 간주함으로써, 노사 응답의 '상대적 위치'가 일관된 정도를 Spearman's rho로 측정할 수 있다. 즉 일치도는 노사 응답의 '절대적 위치'의 일치성을 평가하며, rho는 '상대적 위치'의 일관성을 평가한다.

<표 5-2>는 경영의사결정 참여 설문문항들에 대한 노사 응답의 신뢰도를 제시한다. 이 설문문항들을 명목척도로 가정하고 “해당없음”을 포함한 신뢰도는 항목들에 따라 53~63%의 일치도를 보여준다. “해당없음”을 결측치로 처리하고 서열척도로 추정한 rho는 0.35~0.65로 나타났다. 일치도는 참여에 대한 노사 관점이 상당한 수준에서 일치하고 있음을 보여주지만, rho는 배치 전환, 명예퇴직/정리해고, 임금 및 근로시간제 등 근로자들의 이해관계가 직접적으로 얽힌 영역에서 노사 관점의 차이를 보여준다. 이러한 차이에도 불구하고 조사된 모든 경영의사결정 영역에서 노사 간 일치도가 50% 이상으로 나타난 것은 근로자 대표조직의 경영 참여에 대한 노사의 관점이 2013년에는 상당한 수준에서 일치하였음을

〈표 5-2〉 경영의사결정 참여 설문문항에 대한 노사 응답의 신뢰도(2013년)

설문문항	일치도 <sup>1)</sup>	rho <sup>2)</sup>
PQ5002/RQ3002: 경영 전략	0.57	0.54
PQ5003/RQ3003: 투자 전략	0.59	0.63
PQ5004/RQ3004: 신기술·신기계의 도입	0.56	0.58
PQ5005/RQ3005: 외주 하청	0.61	0.62
PQ5006/RQ3006: 국내외 공장이전	0.62	0.57
PQ5007/RQ3007: 근로자의 배치전환	0.60	0.41
PQ5008/RQ3008: 비정규직 채용	0.53	0.50
PQ5010/RQ3010: 명예퇴직/정리해고	0.56	0.43
PQ5011/RQ3011: 인사평가제도 변경	0.57	0.49
PQ5012/RQ3012: 임금제도 변경	0.62	0.39
PQ5014/RQ3014: 승진제도 변경	0.57	0.49
PQ5015/RQ3015: 근로시간제도 변경	0.63	0.43

주: 1) 표본수는 587개임.

2) “97) 해당없음”이 결측치로 처리되었기 때문에 설문문항에 따라 표본수가 269개에서 519개로 다름.

〈표 5-3〉 경영의사결정 참여 설문문항의 내적 일관도

	2005	2007	2009	2011	2013	2015
코딩방식 1	0.83	0.89	0.90	0.92	0.92	0.90
코딩방식 2	0.79	0.88	0.90	0.91	0.91	0.87

주: 코딩방식 1은 “1) 사전통보도 없이 회사가 결정”, “2) 사전통보”, “97) 해당없음”은 0, “3) 노사협의”와 “4) 노사공동결정”은 1로 코딩; 코딩방식 2는 “1) 사전통보도 없이 회사가 결정”, “2) 사전통보”, “3) 노사협의”, “97) 해당없음”은 0, “4) 노사공동결정”은 1로 코딩.

보여주며, 동시에 다른 연도들의 사측 응답이 특별하게 사측 관점에 편향되었을 것으로 간주할 이유가 없음을 시사한다.

〈표 5-3〉은 총 11개의 경영의사결정 참여 설문문항들에 대한 사측 응답이 얼마나 일관되는지를 가리키는 내적 일관도를 연도별, 코딩방식별로 보여준다. 코딩방식 1의 내적일관도는 0.83부터 0.92, 코딩방식 2의 내적일관도는 0.79부터 0.91로 나타났다. 2005년도의 내적 일관도가 다소 낮게 나타났지만 두 코딩방식 모두 받아들일 수 있는 전통적인 기준보다 높게

나타났다. 따라서 코딩방식 1과 2 모두 경영 참여 수준을 측정할 수 있는 지수를 사용하기에 적합한 것으로 볼 수 있다.

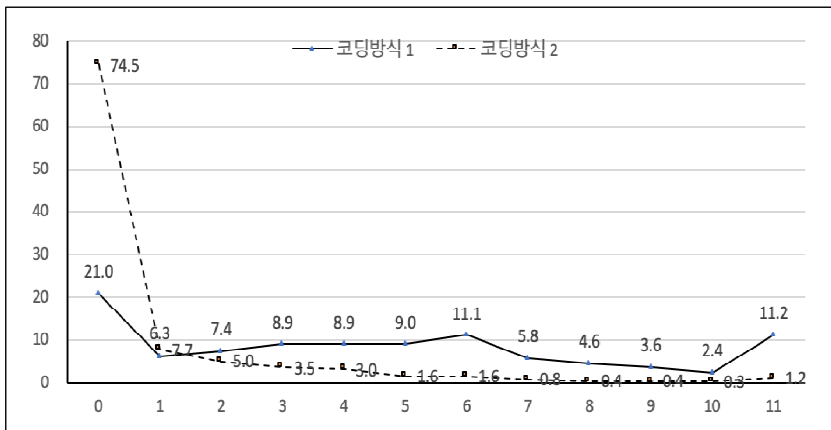
## 2. 경영의사결정 참여수준의 상관관계 분석

### 가. 경영의사결정 참여수준의 기초통계

사업체 간 근로자 대표조직의 경영의사결정 참여 수준을 비교하기 위한 가장 단순하고 직접적인 방법은 총 11개의 설문문항들에 대한 사측 응답들의 단순합을 계산하고 비교하는 것이다. [그림 5-2]는 2005년부터 2015년까지 코딩방식에 따른 사측 응답의 단순합 비율 분포를 보여준다.

“노사협의”와 “공동결정”을 참여로 간주한 코딩방식 1에 따른 단순합의 비율 분포는 참여가 전무한 사업체가 21.0%, 6개 영역의 참여가 11.1%, 11개 영역 모두 참여가 11.2%로 나타났다. 다음으로 “공동결정”만을 참여로 간주한 코딩방식 2에 따른 단순합의 비율 분포는 참여가 전무한 사업체가 74.5%, 6개 영역의 참여가 1.6%, 11개 영역 모두 참여가

[그림 5-2] 경영의사결정 참여의 단순합 분포(2005~2015년)



주: 코딩방식 1은 “(1) 사전통보도 없이 회사가 결정”, “(2) 사전통보”, “(97) 해당없음”은 0, “(3) 노사협의”와 “(4) 노사공동결정”은 1로 코딩; 코딩방식 2는 “(1) 사전통보도 없이 회사가 결정”, “(2) 사전통보”, “(3) 노사협의”, “(97) 해당없음”은 0, “(4) 노사공동결정”은 1로 코딩.



1.2%로 나타났다. 따라서 근로자 대표조직이 경영의사결정에 참여할 경우, 그 참여 방식은 “노사협의”가 절대적 다수임을 알 수 있다. 연도별 단순합의 빈도수와 비율은 <부표 5-1>에 제시되어 있다.

<표 5-4>는 경영참여 정의에 따른 단순합과 Rasch 참여지수의 연도별 평균 및 표준편차를 제시한다. 먼저 단순합 참여지수를 살펴보면 다음과 같다. 참여지수 1은 2005년 3.89에서 2009년 4.95로 약 1개 영역에서 참여가 증가하였으나, 이후 2015년 4.29로 감소한 것으로 나타났다. 참여지수 2는 2013년 0.93을 제외하면 연도별 참여수준의 변화가 크지 않은 것으로 나타났다. 전반적으로 두 지수의 표준편차는 증가하는 것으로 나타났다.

다음으로 Rasch 참여지수를 살펴보면 다음과 같다. 참여지수 1은 2005년 -0.042에서 2009년 1.211로 평균값이 크게 증가하였으나, 이후 특별한 경향성은 나타나지 않았다. 다만 단순합의 증감과 Rasch 추정값의 증감이 반드시 일치하지 않음을 보여주며, 이러한 경향성은 참여지수 2에서도

<표 5-4> 경영의사결정 참여수준의 평균과 표준편차

	표본 수	참여지수 1		참여지수 2	
		단순합	Rasch	단순합	Rasch
2005	1,041	3.890 (2.979)	-0.042 (1.950)	0.858 (1.648)	-2.324 (1.264)
2007	1,219	4.372 (3.517)	0.273 (2.313)	0.852 (1.979)	-2.392 (1.389)
2009	1,300	4.945 (3.657)	1.211 (2.477)	0.849 (2.070)	-2.495 (1.286)
2011	1,412	4.612 (3.831)	0.313 (2.585)	0.837 (2.101)	-2.484 (1.426)
2013	1,270	4.728 (3.822)	0.734 (2.726)	0.928 (2.179)	-2.126 (1.468)
2015	950	4.286 (3.580)	0.813 (2.688)	0.857 (1.940)	-1.981 (1.403)

주: 참여지수 1은 코딩방식 1, 참여지수 2는 코딩방식 2를 기반으로 지수화한 것임.

나타났다.<sup>38)</sup>

#### 나. 경영의사결정 참여수준과 Correlates 분석

<표 5-5>는 2005년부터 2015년까지 드러난 경영의사결정 참여지수와 노사관계의 전략변수 및 성과변수들 간 상관관계를 보여준다.

먼저 참여지수들 간 상관관계는 다음과 같다. 참여지수 1(2)의 단순합과 Rasch 간 상관관계는 0.986(0.975)로 매우 높았으며, 단순합(Rasch) 지수들 간 상관관계는 0.362(0.354)로 나타났다.

<표 5-5> 경영의사결정 참여수준과 Correlates

		참여지수 1		참여지수 2		해당없음 (단순합)	해당없음 (Rasch)
		단순합	Rasch	단순합	Rasch		
참여지수 1	단순합						
	Rasch	0.986 ***					
참여지수 2	단순합	0.362 ***	0.351 ***				
	Rasch	0.356 ***	0.354 ***	0.975 ***			
노사 파트너십	사측응답	0.111 ***	0.101 ***	0.080 ***	0.084 ***	-0.015	-0.001
	노측응답	0.121 ***	0.107 ***	0.072 ***	0.070 ***	0.000	-0.003
협조전략	사측응답	0.009	0.012	0.047 **	0.038 *	0.060 ***	0.057 ***
	노측응답	0.070 ***	0.059 ***	0.064 ***	0.064 ***	0.013	0.000
노사관계 분위기	단일문항 <sup>1)</sup>	0.065 ***	0.067 ***	0.049 ***	0.050 ***	-0.006	0.004
	사측응답	0.021	0.003	0.100 ***	0.105 ***	0.010	0.041 *
	노측응답	0.07 ***	0.066 ***	0.100 ***	0.097 ***	0.028	0.027
고용안정성		0.123 ***	0.122 ***	0.049 ***	0.064 ***	-0.045 ***	-0.051 ***
노동생산성		0.031 *	0.029 *	0.014	0.015	-0.013	-0.010
1인당 부가가치		0.066 ***	0.063 ***	0.047 **	0.050 **	-0.062 ***	-0.060 ***
재무적 성과		0.028 *	0.026 +	0.005	0.004	-0.014	-0.010
1인당 영업이익		0.064 ***	0.066 ***	0.057 ***	0.060 ***	-0.066 ***	-0.064 ***

주: 1) 노사관계분위기를 측정하는 단일문항은 2015년에 조사되지 않았음.

38) 연도별로 경영의사결정 영역별로 가중치를 추정하였기 때문에, Rasch 참여지수를 연도별로 직접적으로 비교하기 어렵다는 것에 유의할 필요가 있다.

경영의사결정 참여 수준에 영향을 주는 선행변수로서 회사와 노동조합의 노사관계정책과 참여지수 간의 상관관계를 살펴본다. 회사의 노사 파트너십 정책에 대한 사측 응답과의 상관관계는 참여지수 1의 단순합(Rasch)이 0.111(0.101), 참여지수 2의 단순합(Rasch)은 0.080(0.084)으로 나타났다. 반면 노측 응답과의 상관관계는 참여지수 1의 단순합(Rasch)이 0.121(0.107)이었으며, 참여지수 2의 단순합(Rasch)이 0.072(0.070)로 나타났다. 다음으로 노동조합의 협조전략에 대한 사측 응답과의 상관관계는 참여지수 1의 단순합(Rasch)이 0.009(0.012), 참여지수 2의 단순합(Rasch)은 0.047(0.038)로 나타났다. 반면 노측 응답과의 상관관계는 참여지수 1의 단순합(Rasch)이 0.070(0.059)이었으며, 참여지수 2의 단순합(Rasch)이 0.064(0.064)로 나타났다. 노동조합의 협조전략은 사측응답과 참여지수 1의 단순합과 Rasch를 제외하고 경영의사결정 참여지수와 통계적으로 유의한 정(+)의 관계에 있었다. 요약하면 회사의 노사 파트너십 정책과 노동조합의 협조전략은 경영의사결정 참여지수와 통계적으로 유의한 정(+)의 관계에 있었다. 특히 사측 인사/노사 담당자의 응답을 이용하여 계산된 참여지수와 노측 응답으로 계산된 노사관계정책들과 상관관계가 더 높다는 것은 적어도 동일 응답자에 의한 편의 가능성이 작다는 점을 시사한다.

경영의사결정 참여 수준에 의해 영향을 받는 노사관계분위기와 생산성 및 재무성과 변수들을 살펴본다. 먼저 전체 표본을 대상으로 단일문항으로 측정된 노사관계분위기와 참여지수 1의 단순합(Rasch) 간 상관관계가 0.065(0.067), 참여지수 2의 단순합(Rasch)은 0.049(0.050)로 나타났다. 한편 사측과 노측 응답을 기반으로 한 노사관계분위기는 노동조합의 노사관계정책과 유사한 상관관계 패턴을 보여준다. 즉 사측 응답의 노사관계분위기와 참여지수 1의 단순합과 Rasch를 제외하고 경영의사결정 참여지수와 통계적으로 유의한 정(+)의 관계에 있었다.

고용안정성과 참여지수 1의 단순합(Rasch) 간 상관관계가 0.123(0.122), 참여지수 2의 단순합(Rasch)은 0.049(0.064)로 나타났으며, 모두 통계적으로 유의한 정(+)의 관계에 있었다.

“동종업체 평균과의 비교”를 기반으로 사측이 응답한 정성적 성과지표

는 정량적 성과지표와 상당히 다른 패턴을 보여준다. 우선 정성적 성과지표를 살펴보면 다음과 같다. 노동생산성과 참여지수 1의 단순합(Rasch) 간 상관관계는 0.031(0.029)이었으며, 재무성과와 참여지수 1의 단순합(Rasch) 간 상관관계는 0.028(0.026)로 나타났으나, 이들만이 통계적으로 유의하였다. 반면 정량적 성과지표로서 1인당 부가가치와 참여지수 1의 단순합(Rasch) 간 상관관계가 0.066(0.063), 참여지수 2의 단순합(Rasch)은 0.047(0.050), 1인당 영업이익과 참여지수 1의 단순합(Rasch) 간 상관관계가 0.064(0.066), 참여지수 2의 단순합(Rasch)은 0.057(0.060)으로 나타났다. 요약하면, 정성적 성과지표보다 정량적 성과지표가 경영의사결정 참여지수와 높은 상관관계를 보여주었다. 이는 사업체패널조사 내 HRM 또는 작업장 참여 변수들이 정성적 성과지표와 높은 상관관계를, 정량적 성과지표와는 낮은 상관관계를 보이는 것과 다른 결과라 할 수 있다.<sup>39)</sup>

마지막으로 “해당없음” 지수를 살펴보면 다음과 같다. 노동조합의 협조 전략에 대한 사측 응답과 “해당없음” 지수의 단순합(Rasch) 간 상관관계는 0.060(0.057), 노사관계분위기의 사측 응답과 “해당없음”지수의 Rasch 간 상관관계가 0.041로 통계적으로 유의한 것으로 나타났다. 한편 고용안정성과 정량적 성과지표인 1인당 부가가치와 1인당 영업이익과는 통계적으로 유의한 수준에서 부(-)의 상관관계를 보여준다. 따라서 경영의사결정 참여를 측정하는 설문문항에서 “97) 해당없음”이라는 응답은 의미있는 정보를 가지고 있지만, 어떤 의미인지는 연구자에게 알려주지 않는다고 할 수 있다.

39) 사업체패널조사를 이용한 논문들이 주로 정성적 성과지표를 분석해 온 경향이 있다는 문제제기와도 관련이 있다(송인수, 2015). 한편 본 연구에서 이러한 결과가 나오게 된 이유로 정량적 성과지표 데이터에 결측치가 많아 표본수가 크게 감소한 것을 고려할 수 있다. 예를 들어, 참여지수와 노동생산성의 상관관계는 표본수 5,084개, 참여지수와 1인당 부가가치의 상관관계는 표본수 3,160개를 사용하여 계산되었다. 만약 표본수의 감소가 무작위로 발생한 것이라면 문제가 없을 것이지만, 재무데이터 등 정량적 성과지표의 가용성(availability)과 근로자 대표조직의 참여 수준이 회사 규모 또는 생산성/수익성과 동시에 관련이 있다면 본 연구와 같이 높은 상관관계가 나타날 수 있다. 따라서 결정요인 분석 시 회사 규모와 다른 특성들을 통제할 필요가 있다.

### 3. 경영의사결정 참여수준의 결정요인 분석

#### 가. 실증모형

이 섹션에서는 경영의사결정 참여수준의 결정요인으로서 사업체 내 노동조합 조직 여부와 노사관계정책들의 영향을 추정한다. 이를 위한 추정식은 다음과 같다.

$$I_{it} = \alpha + \beta U_{it} + \gamma X_{it} + u_i + Y_t + e_{it} \quad (5-3)$$

$I_{it}$ 는 종속변수로 사업체  $i$ 의 연도  $t$ 에서의 참여지수,  $U_{it}$ 는 사업체  $i$ 의 연도  $t$ 에서의 노동조합 조직 여부(또는 노사의 노동관계정책),  $X_{it}$ 는 사업체  $i$ 의 연도  $t$ 에서의 사업체 특성,  $u_i$ 는 관측되지 않은 시간불변의 사업체 특성,  $Y_t$ 는 연도  $t$ 의 고정효과, 그리고  $e_{it}$ 는 사업체  $i$ 의 시간  $t$ 에서의 오차항을 가리킨다.

시간에 따라 변화하는 사업체 특성, 즉  $X_{it}$ 는 참여지수에 영향을 미칠 수 있는 사업체의 규모, 연령, 사업체 근로자들의 특성을 통제하기 위한 여성, 생산직, 단순직, 그리고 정규직 비율, 가치사슬 구조의 영향을 통제하기 위한 하도급거래 여부를 포함하였다.

#### 나. 경영의사결정 참여수준의 결정요인 결과

<표 5-6>은 경영의사결정 참여지수의 결정요인으로서 노동조합과 회사 및 노동조합의 노사관계정책이 미치는 영향을 보여준다. 먼저 전체표본을 대상으로 한 노동조합 조직 유무의 계수를 살펴보면 다음과 같다. 종속변수로서 참여지수 1의 단순합과 Rasch에 대해 각각 1.492( $p < 0.001$ )와 1.030( $p < 0.001$ )으로 나타났으며, 참여지수 2의 단순합과 Rasch에 대해 0.963( $p < 0.001$ )과 0.694( $p < 0.001$ )로 나타났다. 즉 유노조사업체의 참여수준이 무노조사업체의 참여수준보다 통계적으로 유의한 수준에서 높은 것으로 나타났다. 예를 들어, 유노조사업체에서 평균적으로 약 1.5개 영역에서 “노사협의” 또는 “공동결정”이 높았으며, 또는 약 1개 영역에서 “공동

〈표 5-6〉 패널고정효과 모형으로 본 경영의사결정 참여수준의 결정요인

변수	전체 표본				노동조합 표본			
	참여지수 1		참여지수 2		참여지수 1		참여지수 2	
	단순합	Rasch	단순합	Rasch	단순합	Rasch	단순합	Rasch
노동조합	1.492*** [0.300]	1.030*** [0.207]	0.963*** [0.180]	0.694*** [0.122]				
파트너십 (노측응답)					0.402** [0.141]	0.255** [0.093]	0.088 [0.104]	0.056 [0.070]
협조전략 (노측응답)					0.234+ [0.121]	0.160* [0.080]	0.291** [0.089]	0.190** [0.060]
사업체 규모 (/100)	0.036* [0.015]	0.025* [0.010]	0.010 [0.009]	0.008 [0.006]	0.012 [0.018]	0.007 [0.012]	0.012 [0.014]	0.011 [0.009]
사업체 연령	-0.081 [0.071]	-0.053 [0.049]	0.005 [0.042]	0.001 [0.029]	-0.080 [0.083]	-0.048 [0.055]	0.032 [0.061]	0.033 [0.041]
여성 비율	-0.011 [0.007]	-0.008 [0.005]	-0.002 [0.004]	-0.001 [0.003]	0.007 [0.014]	0.007 [0.009]	-0.021+ [0.011]	-0.013+ [0.007]
생산직 비율	-0.004+ [0.003]	-0.003+ [0.002]	-0.002 [0.002]	-0.001 [0.001]	-0.002 [0.003]	-0.001 [0.002]	0.001 [0.002]	0.000 [0.002]
단순직 비율	-0.004 [0.003]	-0.002 [0.002]	-0.001 [0.002]	-0.001 [0.001]	0.000 [0.004]	0.000 [0.003]	0.001 [0.003]	0.000 [0.002]
정규직 비율	0.003 [0.004]	0.003 [0.003]	0.002 [0.002]	0.001 [0.002]	0.008 [0.007]	0.005 [0.004]	0.003 [0.005]	0.002 [0.003]
하도급거래	-0.159 [0.117]	-0.112 [0.081]	-0.100 [0.070]	-0.055 [0.048]	-0.171 [0.169]	-0.123 [0.112]	-0.268* [0.125]	-0.171* [0.084]
_cons	4.773** [1.485]	0.453 [1.022]	0.222 [0.891]	-2.760*** [0.605]	5.290* [2.224]	0.693 [1.476]	-0.230 [1.644]	-3.293** [1.111]
연도 통제	예	예	예	예	예	예	예	예
R <sup>2</sup>	0.024	0.053	0.007	0.030	0.032	0.079	0.010	0.023
Obs	7,192	7,192	7,192	7,192	3,399	3,399	3,399	3,399
Grp_No	1,550	1,550	1,550	1,550	752	752	752	752

주: \*\*\*는 0.1% 수준\*\*는 1% 수준, \*는 5% 수준, +는 10% 수준에서 통계적으로 유의함.

결정”이 높은 것으로 나타났다.

노동조합 표본만을 대상으로 회사 및 노동조합의 노사관계정책이 미치는 영향을 살펴보면 다음과 같다. 종속변수로서 참여지수 1의 단순합에 대해 회사의 파트너십(노측 응답)과 노동조합의 협조전략(노측 응답)은 각각 0.402( $p < 0.01$ )와 0.234( $p < 0.10$ )로, 참여지수 1의 Rasch에 대해 0.255( $p < 0.01$ )와 0.160( $p < 0.05$ )으로 나타났다. 다음으로 종속변수로서 참여지수 2의 단순합에 대해 회사의 파트너십(노측 응답)과 노동조합의 협조전략(노측 응답)은 각각 0.088( $p > 0.10$ )과 0.291( $p < 0.01$ )로, 참여지수 2의 Rasch에 대해 0.056( $p > 0.10$ )과 0.190( $p < 0.01$ )으로 나타났다. 요약하면 참여지수 1과 관련하여 회사의 파트너십과 노동조합의 협조전략이 모두 통계적으로 유의하였으나, 참여지수 2와 관련하여 노동조합의 협조전략만이 통계적으로 유의하였다.

표에는 제시하지는 않았지만, 회사의 파트너십(사측 응답)과 노동조합의 협조전략(사측 응답) 결과는 다음과 같다. 먼저 종속변수로 참여지수 1의 단순합에 대해 0.410( $p < 0.01$ )과 -0.085( $p > 0.10$ )로, 참여지수 1의 Rasch에 대해 0.253( $p < 0.01$ )과 -0.041( $p > 0.10$ )로 나타났다. 참여지수 2의 단순합에 대해 0.104( $p > 0.10$ )과 0.095( $p > 0.10$ )로, 참여지수 2의 Rasch에 대해 0.207( $p < 0.05$ )과 0.113( $p < 0.10$ )으로 나타났다.

#### 4. 경영의사결정 참여수준과 사업체 성과 분석

##### 가. 실증모형

이 섹션에서는 근로자 대표조직, 즉 노동조합과 노사협의회에 통한 경영의사결정 참여수준이 사업체의 노사관계 및 경제적 성과에 미친 영향을 추정한다. 이를 위한 추정식은 다음과 같다.

$$y_{it} = \alpha + \beta I_{it} + \gamma X_{it} + u_i + Y_t + e_{it} \quad (5-4)$$

$y_{it}$ 는 종속변수로 사업체  $i$ 의 연도  $t$ 에서의 노사관계분위기, 생산성, 고용안정성, 수익성이고,  $I_{it}$ 는 사업체  $i$ 의 연도  $t$ 에서의 경영의사결정 참

여 수준,  $X_{it}$ 는 사업체  $i$ 의 연도  $t$ 에서의 사업체 특성,  $u_i$ 는 관측되지 않은 시간불변의 사업체 특성,  $Y_t$ 는 연도  $t$ 의 고정효과, 그리고  $e_{it}$ 는 사업체  $i$ 의 시간  $t$ 에서의 오차항을 가리킨다.

시간에 따라 변화하는 사업체 특성, 즉  $X_{it}$ 는 노사관계 및 경제적 성과에 영향을 미칠 수 있는 사업체의 규모, 연령, 사업체 근로자들의 특성을 통제하기 위한 여성, 생산직, 단순직, 그리고 정규직 비율, 가치사슬 구조의 영향을 통제하기 위한 하도급거래 여부를 포함하였다.

#### 나. 경영의사결정 참여수준과 노사관계분위기

<표 5-7>은 경영의사결정 참여지수의 노사관계분위기에 대한 패널고정효과 결과를 보여준다. 먼저 단일문항으로 측정된 노사관계분위기 결과는 다음과 같다. 참여지수 1의 단순합과 Rasch 계수는 각각 0.006( $p<0.05$ )과 0.010( $p<0.05$ )으로 나타났으며, 참여지수 2의 단순합과 Rasch 계수는 각각 0.013( $p<0.01$ )과 0.020( $p<0.01$ )으로 나타났다.

다음으로 복합문항으로 측정된 노측 응답의 노사관계분위기 결과는 다음과 같다. 참여지수 1의 단순합과 Rasch 계수는 각각 0.004( $p>0.10$ )와 0.005( $p>0.10$ )로 통계적으로 유의하지 않은 반면 참여지수 2의 단순합과 Rasch 계수는 각각 0.023( $p<0.001$ )과 0.032( $p<0.001$ )로 통계적으로 유의하였다.

사측 응답의 노사관계분위기 결과를 제시하지는 않았지만, 참여지수 1의 단순합과 Rasch 계수는 각각 -0.006( $p<0.10$ )과 -0.010( $p<0.05$ )으로 통계적으로 유의한 음(-)의 관계로 나타났다. 반면 참여지수 2의 단순합과 Rasch 계수는 각각 0.016( $p<0.001$ )과 0.025( $p<0.001$ )로 통계적으로 유의하였다.

요약하면 참여지수 1과는 다르게, 참여지수 2의 단순합과 Rasch는 노사관계분위기와 관련하여 일관된 결과를 보여주었다. 또한 단순합과 Rasch의 계수를 비교하면 Rasch의 계수가 단순합보다 일관되게 더 크게 추정되었다.



〈표 5-7〉 패널고정효과 모형으로 본 참여수준의 노사관계분위기에 대한 영향

변수	노사관계분위기 (단일문항)				노사관계분위기 (노측 응답)			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
사업체 규모(/100)	0.002 [0.003]	0.002 [0.003]	0.002 [0.003]	0.002 [0.003]	-0.001 [0.003]	-0.001 [0.003]	-0.001 [0.003]	-0.001 [0.003]
사업체 연령	-0.018*** [0.003]	-0.019*** [0.003]	-0.018*** [0.003]	-0.018*** [0.003]	0.013 [0.015]	0.013 [0.015]	0.012 [0.015]	0.011 [0.015]
여성 비율	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]	-0.003 [0.003]	-0.003 [0.003]	-0.003 [0.003]	-0.003 [0.003]
생산직 비율	0.001 [0.000]	0.001 [0.000]	0.001 [0.000]	0.001 [0.000]	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]
단순직 비율	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]	-0.001 [0.001]	-0.001 [0.001]	-0.001 [0.001]	-0.001 [0.001]
정규직 비율	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]
하도급 거래	0.000 [0.023]	0.000 [0.023]	0.001 [0.023]	0.001 [0.023]	-0.028 [0.031]	-0.028 [0.031]	-0.022 [0.031]	-0.023 [0.031]
참여지수 1 (단순합)	0.006* [0.003]				0.004 [0.004]			
참여지수 1 (Rasch)		0.010* [0.004]				0.005 [0.005]		
참여지수 2 (단순합)			0.013** [0.004]				0.023*** [0.005]	
참여지수 2 (Rasch)				0.020*** [0.006]				0.032*** [0.007]
_cons	4.177*** [0.116]	4.208*** [0.116]	4.178*** [0.116]	4.245*** [0.117]	3.314*** [0.412]	3.331*** [0.411]	3.329*** [0.409]	3.429*** [0.410]
연도 통제	예	예	예	예	예	예	예	예
R <sup>2</sup>	0.014	0.014	0.014	0.015	0.015	0.015	0.023	0.022
Obs	5,911	5,911	5,911	5,911	3,272	3,272	3,272	3,272
Grp_No	1,389	1,389	1,389	1,389	671	671	671	671

주: \*\*\*는 0.1% 수준\*\*는 1% 수준, \*는 5% 수준, +는 10% 수준에서 통계적으로 유의함.

## 다. 경영의사결정 참여수준과 고용안정성

<표 5-8>은 패널고정효과 모형의 로짓모형을 이용하여 추정한 경영의

<표 5-8> 패널고정효과 모형으로 본 참여수준의 고용안정성에 대한 영향

	(1)	(2)	(3)	(4)
사업체규모(/100)	0.024 <sup>+</sup> [0.014]	0.024 <sup>+</sup> [0.014]	0.026 <sup>+</sup> [0.014]	0.026 <sup>+</sup> [0.014]
사업체 연령	-0.467 [16.581]	-0.502 [27.470]	-0.5 [24.676]	-0.494 [22.324]
여성 비율	-0.001 [0.007]	-0.001 [0.007]	-0.002 [0.007]	-0.002 [0.007]
생산직 비율	-0.001 [0.002]	-0.001 [0.002]	-0.002 [0.002]	-0.002 [0.002]
단순직 비율	-0.002 [0.003]	-0.002 [0.003]	-0.002 [0.003]	-0.002 [0.003]
정규직 비율	0.001 [0.003]	0.001 [0.003]	0.001 [0.003]	0.001 [0.003]
하도급거래	0.006 [0.100]	0.006 [0.100]	-0.001 [0.100]	-0.001 [0.100]
참여지수 1 (단순합)	0.046 <sup>***</sup> [0.011]			
참여지수 1 (Rasch)		0.069 <sup>***</sup> [0.017]		
참여지수 2 (단순합)			0.005 [0.018]	
참여지수 2 (Rasch)				0.016 [0.026]
연도통제	예	예	예	예
ll	-3231.381	-3230.884	-3261.630	-3259.851
Obs	3,891	3,891	3,891	3,891
Grp_No	783	783	783	783

주: \*\*\*는 0.1% 수준\*\*, \*는 1% 수준, +는 5% 수준, +는 10% 수준에서 통계적으로 유의함.

사결정 참여의 고용안정성에 대한 영향을 보여준다. 고용안정성과 관련하여 참여지수 1의 단순합과 Rasch 계수는 각각 0.046( $p<0.001$ )과 0.069( $p<0.001$ )로 나타난 반면 참여지수 2의 단순합과 Rasch 계수는 각각 0.005( $p>0.10$ )와 0.016( $p>0.10$ )으로 나타났다.

다른 성과변수들과는 다르게, 참여지수 1이 통계적으로 유의한 반면 참여지수 2는 그렇지 않았다. 이러한 결과는 “노사협의” 방식의 참여로도 고용안정성을 높일 수 있음을 의미한다. 또한 고용안정성이 시간상 변화가 없었던 767개 사업체의 총 3,301개 표본이 분석에서 제외된 것 또한 참여지수 2가 통계적으로 유의하지 않게 된 이유로도 생각된다.

#### 라. 생산성과 수익성

다음으로 경영의사결정 참여 수준에 따른 사업체의 경제적 성과를 살펴본다. <표 5-9>의 생산성 추정결과를 살펴보면, 정성적 지표로서 사측이 응답한 “동종업체 평균”과 비교한 사업체의 노동생산성과 관련하여 경영참여 지수 1의 단순합과 Rasch 계수가 각각 0.005( $p<0.10$ )와 0.007( $p<0.10$ )로 나타났으며, 참여지수 2의 단순합과 Rasch 계수는 각각 0.009( $p<0.10$ )와 0.014( $p<0.05$ )로 나타났다. 반면 정량적 지표로서 1인당 부가 가치와 관련해서는 어떠한 참여지수도 통계적으로 유의하지 않았다.

마지막으로 참여지수와 수익성 간 관계를 살펴보면 <표 5-10>과 같다. 수익성을 나타내는 정성적 지표로서 “동종업체 평균” 대비 사업체의 재무적 성과와 관련하여 참여지수 1의 단순합과 Rasch 계수는 각각 0.003( $p>0.10$ )과 0.004( $p>0.10$ )로 나타난 반면 참여지수 2의 단순합과 Rasch 계수는 각각 0.012( $p<0.05$ )와 0.021( $p<0.01$ )로 나타났다. 반면 수익성의 정량적 지표로서 1인당 영업이익과 관련하여 어떠한 참여지수도 통계적으로 유의하지 않았다. 이러한 결과의 패턴은 앞서 살펴본 1인당 부가 가치와도 동일한 것이었다.

정성적 성과지표와의 결과를 요약하면, 참여지수 1은 노동생산성과 통계적으로 유의하게 나타났으나 재무성과와는 유의하지 않았다. 반면 참여지수 2는 두 변수 모두와 통계적으로 유의하게 나타났다. 단순합과 Rasch

〈표 5-9〉 패널고정효과 모형으로 본 참여수준의 노동생산성에 대한 영향

	노동생산성				1인당 부가가치			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
사업체 규모(/100)	-0.007* [0.003]	-0.007* [0.003]	-0.007* [0.003]	-0.007* [0.003]	-0.009** [0.003]	-0.009** [0.003]	-0.009** [0.003]	-0.009** [0.003]
사업체 연령	-0.008* [0.003]	-0.008* [0.003]	-0.008* [0.003]	-0.008* [0.003]	-0.017 [0.019]	-0.018 [0.019]	-0.018 [0.019]	-0.018 [0.019]
여성 비율	-0.001 [0.002]	-0.001 [0.002]	-0.002 [0.002]	-0.002 [0.002]	-0.004* [0.002]	-0.004* [0.002]	-0.004* [0.002]	-0.004* [0.002]
생산직 비율	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]
단순직 비율	0.001* [0.001]	0.001* [0.001]	0.001* [0.001]	0.001* [0.001]	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]
정규직 비율	-0.002* [0.001]	-0.002* [0.001]	-0.002* [0.001]	-0.002* [0.001]	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]
하도급 거래	0.023 [0.024]	0.023 [0.024]	0.023 [0.024]	0.023 [0.024]	0.008 [0.024]	0.008 [0.024]	0.007 [0.024]	0.007 [0.024]
참여지수 1 (단순합)	0.005+ [0.003]				0.003 [0.003]			
참여지수 1 (Rasch)		0.007+ [0.004]				0.004 [0.004]		
참여지수 2 (단순합)			0.009+ [0.005]				0.000 [0.005]	
참여지수 2 (Rasch)				0.014* [0.007]				0.001 [0.007]
_cons	3.568*** [0.128]	3.592*** [0.129]	3.572*** [0.128]	3.617*** [0.130]	4.637*** [0.411]	4.649*** [0.411]	4.646*** [0.411]	4.648*** [0.412]
연도 통제	예	예	예	예	예	예	예	예
R <sup>2</sup>	0.007	0.007	0.007	0.008	0.046	0.046	0.046	0.046
Obs	5,084	5,084	5,084	5,084	3,140	3,140	3,140	3,140
Grp_No	1,229	1,229	1,229	1,229				

주: \*\*\*는 0.1% 수준, \*\*는 1% 수준, \*는 5% 수준, +는 10% 수준에서 통계적으로 유의함.

〈표 5-10〉 패널고정효과 모형으로 본 참여수준의 수익성에 대한 영향

	채무성과				1인당 영업이익			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
사업체 규모(/100)	0.001 [0.003]	0.001 [0.003]	0.001 [0.003]	0.001 [0.003]	-0.008 [0.006]	-0.008 [0.006]	-0.008 [0.006]	-0.008 [0.006]
사업체 연령	-0.012** [0.004]	-0.012** [0.004]	-0.012** [0.004]	-0.012** [0.004]	-0.023 [0.037]	-0.023 [0.037]	-0.022 [0.037]	-0.021 [0.037]
여성 비율	-0.002 [0.002]	-0.002 [0.002]	-0.002 [0.002]	-0.002 [0.002]	-0.005+ [0.003]	-0.005+ [0.003]	-0.005+ [0.003]	-0.005+ [0.003]
생산직 비율	0.001 [0.001]	0.001 [0.001]	0.001 [0.001]	0.001 [0.001]	-0.002 [0.001]	-0.002 [0.001]	-0.002+ [0.001]	-0.002+ [0.001]
단순직 비율	0.001 [0.001]	0.001 [0.001]	0.001 [0.001]	0.001 [0.001]	-0.002 [0.001]	-0.002 [0.001]	-0.002 [0.001]	-0.002 [0.001]
정규직 비율	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]	0.000 [0.001]	-0.001 [0.002]	-0.001 [0.002]	-0.001 [0.002]	-0.001 [0.002]
하도급 거래	-0.021 [0.027]	-0.021 [0.027]	-0.02 [0.027]	-0.02 [0.027]	-0.018 [0.042]	-0.018 [0.042]	-0.020 [0.042]	-0.020 [0.042]
참여지수 1 (단순합)	0.003 [0.003]				0.003 [0.005]			
참여지수 1 (Rasch)		0.004 [0.005]				0.006 [0.008]		
참여지수 2 (단순합)			0.012* [0.005]				-0.009 [0.009]	
참여지수 2 (Rasch)				0.021** [0.008]				-0.018 [0.013]
_cons	3.482*** [0.142]	3.494*** [0.142]	3.478*** [0.142]	3.545*** [0.144]	3.760*** [0.770]	3.774*** [0.770]	3.754*** [0.770]	3.694*** [0.772]
연도 통제	예	예	예	예	예	예	예	예
R <sup>2</sup>	0.004	0.004	0.005	0.005	0.014	0.014	0.015	0.015
Obs	5,110	5,110	5,110	5,110	3,617	3,617	3,617	3,617
Grp_No	1,236	1,236	1,236	1,236	848	848	848	848

주: \*\*\*는 0.1% 수준 \*\*는 1% 수준, \*는 5% 수준, +는 10% 수준에서 통계적으로 유의함.

결과를 요약하면, 단순합보다 Rasch가 통계적으로도 유의한 수준에서 일관되게 큰 추정값을 보여준다.

정량적 지표와의 결과를 요약하면, 참여지수 1과 참여지수 2 모두 통계적으로 유의하지 않았다. 상관관계 분석에서 확인되었던 유의한 상관관계는 연도고정효과에 의해 사라졌음을 확인하였다.

## 제5절 소 결

본 연구는 사업체패널조사의 경영의사결정참여 설문문항들의 효과적인 활용방법을 탐색하기 위하여 다음과 같은 분석을 수행하였다. 첫째, 경영 참여를 측정한 정성적 설문 문항들의 신뢰도를 검증하였다. 둘째, 경영 참여의 타당도를 선행변수들과 성과변수들 간 상관관계로 검증하였다. 셋째, 패널고정효과모형을 통해 경영 참여와 이들 변수들의 타당성을 검토하였다. 그 결과는 다음과 같이 요약할 수 있다.

첫째, 경영의사결정 참여 설문문항들의 신뢰도 검증결과는 다음과 같다. 노사 양측에 모두 조사된 2013년 설문문항들을 분석하여 “해당없음”을 포함한 노사 간 응답들의 일치도가 11개 영역들에서 53~63%로 나타났다. 또한 “해당없음”을 결측치로 처리한 후 추정된 Spearman's rho는 0.35~0.65로 나타났다. 이를 통해 2103년에 한하지만 노사 양측이 사업체 내 참여와 관련하여 상당히 높은 수준에서 일치된 평가를 하고 있음을 알려준다. 이에 따라 사측만이 응답한 다른 연도들의 설문문항들이 특별히 사측 관점에 편향되었을 것으로 볼 이유는 없게 된다. 또한 “노사협의”와 “공동결정”을 광의의 참여(코딩방식 1)로, 그리고 “공동결정”을 협의의 참여(코딩방식 2)로 정의한 후 추정한 사측 응답의 내적 일관도는 연도별로 코딩방식 1이 0.83부터 0.92, 코딩방식 2가 0.79부터 0.91로 나타나 경영 참여 수준을 측정할 수 있는 지수를 사용하기에 적합한 것으로 나타났다.

둘째, 두 가지 방식으로 정의된 경영 참여를 기반으로 사업체 내 근로

자 대표조직의 경영 참여 수준을 측정하기 위하여 단순합과 Rasch 추정값을 참여지수로 사용하였다. 먼저 참여지수 간 상관관계를 요약하면, 동일한 참여지수 내 상관관계는 약 0.98, 참여지수 간 상관관계는 약 0.35~0.36으로 나타났다. 다르게 정의된 참여 간 상관관계가 낮았으며, 이는 협의의 참여의 빈도가 상대적으로 낮기 때문이다.

경영 참여 수준과 노사관계정책들 간 상관관계를 요약하면, 참여지수 1과 노사 파트너십과 협조전략의 상관관계는 0.10~0.12, 참여지수 2는 0.07~0.08로 나타났다. 특히 사측 응답으로 계산된 참여지수와 노측 응답으로 계산된 이 노사관계정책들 간 상관관계가 더 높게 나타나 적어도 동일 응답자에 의한 편의는 없는 것으로 보인다.

경영 참여 수준과 성과지표들 간 상관관계를 요약하면, 정성적 성과지표보다 정량적 성과지표에서 참여지수가 높은 상관관계를 보여주었으며, 이는 사업체패널조사 내 HRM 또는 작업장 참여 변수들이 정성적 성과지표와 높은 상관관계를, 정량적 성과지표와는 낮은 상관관계를 보이는 것과 다른 결과라 할 수 있다

마지막으로, 근로자 대표조직의 경영 참여 수준과 선행변수들, 그리고 성과변수들과의 관계를 패널고정모형을 이용하여 분석한 결과는 다음과 같다. 첫째, 노동조합의 조직 여부가 경영 참여 수준의 중요한 결정요인으로 나타났다. 다음으로 참여지수 1의 결정요인으로 회사의 파트너십과 노동조합의 협조전략이 모두 통계적으로 유의하였으나, 참여지수 2의 결정요인으로서는 협조전략만이 통계적으로 유의하였다. 후자의 결과와 관련하여 노동조합의 협조가 전제되어야 공동결정의 참여가 가능하다는 것을 시사한다.

다음으로, 사업체 성과의 결정요인으로 경영 참여 수준의 영향을 요약하면 다음과 같다. 노사관계분위기와 관련하여 참여지수 2가 참여지수 1보다 통계적으로 유의한 영향을 미치는 것으로 나타났으며, 이는 노사 양측 모두에게서 일관되게 나타났다. 한편 참여지수 1이 높아질수록 고용안정성 또한 통계적으로 유의한 수준에서 높아지는 것으로 나타났으나 참여지수 2는 그렇지 않았다.

정성적 성과지표와 관련하여, 참여지수 1은 노동생산성을 통계적으로

유의한 수준에서 높이는 것으로 나타났으나 재무성과에서는 그렇지 않았다. 참여지수 2는 두 정성적 성과지표들과 통계적으로 유의하게 나타났다. 반면 참여지수 1과 참여지수 2 모두 정량적 성과지표와 통계적으로 유의하지 않은 것으로 나타났다. 이는 상관관계 분석결과와는 다르지만 사업체패널조사를 분석한 국내 연구들의 결과와는 유사한 것으로 볼 수 있다.

위 결과를 요약하자면, 협의의 경영 참여가 광의의 경영 참여보다, 그리고 Rasch 참여지수가 단순합 참여지수보다 결정요인 분석에서 일관되고 통계적으로 유의한 것으로 나타났다.

지금까지 경영의사결정 참여 설문문항들의 요인구조와 관련한 논의를 진행한 것은 아니다. 이는 본 연구의 목적이 단일 요인으로서 경영 참여의 신뢰도와 타당도 검증에 있기 때문이었으며, 동시에 본 연구의 경영 참여 측정이 단일요인으로서 참여지수를 구성하기에 충분한 신뢰도를 가진 것으로 검증되었기 때문이다. 그럼에도 불구하고, 후속 연구가 경영 참여의 하위요인 구조를 뒷받침할 이론적 논의와 실증분석을 진행할 필요가 있다. 저자가 판단하기에 가장 유망한 경로는 전략적 선택이론을 기반으로 사업체 내 경영의사결정 참여를 전략적 수준과 기능적 수준으로 구분하고, 하위요인 구조가 수준별 참여들을 반영하는가를 검증하는 것이다. 불충분하지만 김현동과 이동진(2010)은 액면(face) 또는 내용(content) 타당도를 기반으로 전략적 수준에서 경영 참여를 분석하였다. 이러한 시도가 중요한 이유는 근로자 대표조직의 경영 참여가 작업장 혁신이나 참여적 작업관행의 근로자 직접 참여는 물론, 최근 관심이 높아진 대안적 참여방식과 밀접히 연결되어 있기 때문이다. 따라서, 경영참여변수의 활용도를 높이기 위하여 연구자들이 이 변수를 다루는 방식을 합의할 필요가 있다. 이러한 측면에서 본 연구는 단일 요인으로서 경영 참여가 회사와 노동조합의 노사관계정책, 노사관계분위기, 생산성과 수익성에 미치는 영향을 실증함으로써 그러한 합의를 도출하기 위한 시도라 할 수 있다.

다음으로, 사업체패널조사의 시계열 정보를 제한적으로만 이용하는 횡단면 회귀분석 연구가 여전히 많은 상황에서, 본 연구가 패널고정효과모형을 하나의 분석사례로서 활용했음을 다시 언급하고자 한다. 연구자의



연구목적에 따라 데이터 활용방식이 다른 것은 당연하지만, 사업체패널조사의 정보를 최대한 활용하는 분석기법들, 예를 들어 동적패널분석(dynamic panel modeling)과 잠재성장모형(latent growth modeling) 등을 활용할 수 있는 연구 질문들을 발전시키고 실증할 필요가 있다. 무노조사업체에서 노사협의회 경영 참여의 성장패턴과 그 결정요인을 파악하는 연구를 예로 들 수 있겠다.

마지막으로, 경영의사결정 참여 설문문항과 관련된 한계를 언급할 필요가 있겠다. 사업체패널조사는 유노조사업장의 경우 노동조합 관계자로부터, 무노조사업장은 노사협의회 관계자들을 대상으로 참여 설문문항들을 측정하고 있다. 즉 유노조사업장에서 노사협의회 경영 참여는 측정되지 않고 있는 것이다. 이로 인해 근로자 대표조직인 노동조합과 노사협의회 경영 참여보다 노동조합 유무에 따른 경영 참여와 분석적으로 구분하기 어려운 한계가 존재한다.

또한 노동조합의 경영 참여와 관련하여 응답자가 설문문항들을 어떻게 해석하고 응답하는지 파악할 필요가 있다. 단체협상 타결은 협상의제와 관련하여 노사가 “합의”한 것을 의미한다. 특히 기능적 수준에서, 예를 들어 임금제도 변경과 관련한 노사합의에 대해 응답자가 “노사협의”로 응답할지 “공동결정”으로 응답할지 모호한 점이 있다. 사업체패널조사에서는 상당수의 노동조합 표본에서 “노사협의”로 응답이 이루어졌다는 점에서 그 의미를 명확히 할 필요가 있을 것으로 생각한다.

〈부표 5-1〉 연도별 경영의사결정 참여의 단순합 분포

연도	0	1	2	3	4	5	6	7	8	9	10	11	합 계
코딩방식 1의 단순합													
2005	183 (17.6)	90 ( 8.7)	105 (10.1)	111 (10.7)	120 (11.5)	158 (15.2)	74 ( 7.1)	70 (6.7)	48 (4.6)	27 (2.6)	19 (1.8)	36 (3.5)	1,041 (100.0)
2007	250 (20.5)	77 ( 6.3)	110 ( 9.0)	109 ( 8.9)	116 ( 9.5)	115 ( 9.4)	120 ( 9.8)	76 (6.2)	63 (5.2)	40 (3.3)	23 (1.9)	120 (9.8)	1,219 (100.0)
2009	233 (17.9)	62 ( 4.8)	61 ( 4.7)	167 (12.9)	119 ( 9.2)	105 ( 8.1)	120 ( 9.2)	96 (7.4)	66 (5.1)	61 (4.7)	37 (2.9)	173 (13.3)	1,300 (100.0)
2011	345 (24.4)	70 ( 5.0)	98 ( 6.9)	92 ( 6.5)	114 ( 8.1)	111 ( 7.9)	176 (12.5)	62 (4.4)	62 (4.4)	48 (3.4)	33 (2.3)	201 (14.2)	1,412 (100.0)
2013	298 (23.5)	65 ( 5.1)	81 ( 6.4)	79 ( 6.2)	100 ( 7.9)	95 ( 7.5)	173 (13.6)	62 (4.9)	47 (3.7)	58 (4.6)	37 (2.9)	175 (13.8)	1,270 (100.0)
2015	200 (21.1)	90 ( 9.5)	76 ( 8.0)	82 ( 8.6)	71 ( 7.5)	64 ( 6.7)	132 (13.9)	50 (5.3)	42 (4.4)	25 (2.6)	21 (2.2)	97 (10.2)	950 (100.0)
합 계	1,509 (21.0)	454 ( 6.3)	531 ( 7.4)	640 ( 8.9)	640 ( 8.9)	648 ( 9.0)	795 (11.1)	416 (5.8)	328 (4.6)	259 (3.6)	170 (2.4)	802 (11.2)	7,192 (100.0)
코딩방식 2의 단순합													
2005	696 (66.9)	128 (12.3)	84 ( 8.1)	45 ( 4.3)	43 ( 4.1)	20 ( 1.9)	10 ( 1.0)	7 (0.7)	1 (0.1)	2 (0.2)	0 (0.0)	5 ( 0.5)	1,041 (100.0)
2007	909 (74.6)	92 ( 7.6)	62 ( 5.1)	50 ( 4.1)	30 ( 2.5)	22 ( 1.8)	15 ( 1.2)	13 (1.1)	5 (0.4)	3 (0.3)	4 (0.3)	14 ( 1.2)	1,219 (100.0)
2009	1,009 (77.6)	79 ( 6.1)	45 ( 3.5)	36 ( 2.8)	38 ( 2.9)	21 ( 1.6)	26 ( 2.0)	14 (1.1)	6 (0.5)	8 (0.6)	2 (0.2)	16 ( 1.2)	1,300 (100.0)
2011	1,094 (77.5)	88 ( 6.2)	65 ( 4.6)	36 ( 2.6)	39 ( 2.8)	18 ( 1.3)	20 ( 1.4)	9 (0.6)	10 (0.7)	7 (0.5)	2 (0.1)	24 ( 1.7)	1,412 (100.0)
2013	953 (75.0)	91 ( 7.2)	56 ( 4.4)	41 ( 3.2)	34 ( 2.7)	24 ( 1.9)	22 ( 1.7)	9 (0.7)	4 (0.3)	7 (0.6)	13 (1.0)	16 ( 1.3)	1,270 (100.0)
2015	700 (73.7)	75 ( 7.9)	48 ( 5.1)	42 ( 4.4)	32 ( 3.4)	8 ( 0.8)	23 ( 2.4)	4 (0.4)	3 (0.3)	3 (0.3)	0 (0.0)	12 ( 1.3)	950 (100.0)
합 계	5,361 (74.5)	553 ( 7.7)	360 ( 5.0)	250 ( 3.5)	216 ( 3.0)	113 ( 1.6)	116 ( 1.6)	56 (0.8)	29 (0.4)	30 (0.4)	21 (0.3)	87 ( 1.2)	7,192 (100.0)

## 참고문헌

- 김정우(2018), 「경영 참여가 사업체의 경제적 성과에 미치는 영향에 관한 패널분석」, 『산업노동연구』 24(1), pp.261~295.
- 김현동 · 이동진(2010), 「근로자대표기구의 전략적 참여와 참여적 작업조직활동 모색: 노사관계의 역할을 중심으로」, 『노동정책연구』 10(4), pp.153~179.
- 김수진 · 김기민(2007), 「2006년 사업체패널조사의 조사 설계와 실사 결과」, 『노동리뷰』 26, pp.57~71.
- 박원우 · 김미숙 · 정상명 · 허규만(2007), 「동일방법편(Common Method Bias)의 원인과 해결방안」, 『인사 · 조직연구』 15(1), pp.89~133.
- 박인호 · 김기민(2016), 「2016 사업체패널조사의 추가 표본설계 및 표본추출」, 『품질연구보고서』, pp.3~44.
- 송민수 · 김정우(2016), 「사업체패널 조사 활용 연구의 현황과 과제」, 『산업관계연구』 26(4), pp.1~30.
- 유병홍 · 김동주(2011), 「무노조사업장에서 노사협의회를 통한 간접참여와 근로자 직접참여가 기업성과에 미치는 영향」, 『기업경영연구』 18(3), pp.219~242.
- 이기재(2015), 「제5차 사업체패널 가중치 연구 최종보고서」, 한국조사연구학회.
- 이주희 · 이승협(2004), 『경영 참여의 실태와 과제』, 한국노동연구원.
- 홍민기 · 박인호 · 김기민 · 송주원 · 박경원 · 권혁진(2016), 『패널자료 품질 개선 연구 VI』, 한국노동연구원.
- Brick, J. M., Morganstein, D. R. and C. L. Wolter(1987), "Additional Uses for Keyfitz Selection," Proceedings of the American Statistical Association, Section on Survey Research Methods,

pp.783~791.

Budd, John. W.(2004), "Non-wage forms of compensation," *Journal of Labor Research* 25(4), pp.597-622.

Butani, S., Robertson, K. W., Mueller, K.(1998), "Assigning Permanent Random Numbers to the Bureau of Labor Statistics Longitudinal (Universe) Data Base," Proceedings of the Section on Survey Research Methods, American Statistical Association.

Freeman, Richard B. and James Medoff(1984), *What do unions do*, New York: Basic Books.

Hardouin, Jean-Benoit(2007), "Rasch analysis: Estimation and tests with raschtest," *The Stata Journal* 7(1), pp.22~44.

Kerlinger, Fred N.(1964), *Foundations of Behavioral Research*, Holt, Rinehart and Winston, Inc., New York.

Kochan, Thomas A., Harry C. Katz, and Robert B. Mckersie(1986), *The transformation of American industrial relations*, New York: Basic Books.

Siemsen, E., Roth, A., and P. Oliveira(2010), "Common method bias in regression models with linear, quadratic, and intraction effects," *Organizational Research Methods* 13, pp.456~476.

Skinner, C. J. and J. N. K. Rao(1996), "Estimation in Dual Frame Surveys with Complex Designs," *Journal of the Statistical Association* 91, pp.349~356.

◆ 執筆陣

- 김유빈(한국노동연구원 패널데이터연구실장)
- 박인호(부경대학교 부교수)
- 이지은(한국노동연구원 전문위원)
- 신선옥(한국노동연구원 책임연구원)
- 이해정(한국노동연구원 전문위원)
- 정현상(한국노동연구원 책임연구원)
- 최세림(한국노동연구원 부연구위원)
- 박경원(한양대학교 조교수)

패널자료 품질개선 연구(VIII)

▪ 발행연월일	2018년 12월 24일 인쇄 2018년 12월 28일 발행
▪ 발 행 인	배 규 식
▪ 발 행 처	<b>한국노동연구원</b> 30147 세종특별자치시 시청대로 370 세종국책연구단지 경제정책동 ☎ 대표 (044) 287-6080 Fax (044) 287-6089
▪ 조판 · 인쇄	거목정보산업(주) (044) 863-6566
▪ 등 록 일 자	1988년 9월 13일
▪ 등 록 번 호	제13-155호

© 한국노동연구원 2018      정가 7,000원

ISBN 978-11-260-0236-8