

2010년 인구주택총조사 방법론 연구

A Methodological Study of the
Population and Housing Census of 2010

2007년 11월 23일

통 계 청

제 출 문

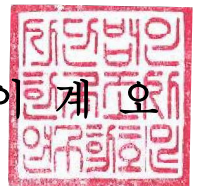
제 출 문

통 계 청 장 귀 하

본 보고서를 “2010년 인구주택총조사 방법론 연구” 과제의 연구결과보고서로 제출합니다.

2007년 11월 23일

한국조사연구학회 이 계 오



연구진

연구책임자 이 건(서울시립대학교 도시사회학과 교수)
공동연구자 강영옥(이화여자대학교 사회생활학과 교수)
전명식(고려대학교 통계학과 교수)
이석훈(충남대학교 통계학과 교수)
김규성(서울시립대학교 통계학과 교수)
이명진(고려대학교 사회학과 교수)
서우석(서울시립대학교 도시사회학과 교수)
연구보조원 남궁상(서울시립대학교 도시사회학과 석사과정)
신지은(충남대학교 통계학과 박사과정)
장중혁(충남대학교 통계학과 석사과정)

연구결과보고서 요약문

연구과제명	2010년 인구주택총조사 방법론 연구		
중심단어	인구주택총조사, 센서스, 조사방법론, 행정자료, 순환조사		
연구기관	한국조사연구학회	연구책임자	이건
연구기간	2007. 3. 27 ~ 2007. 11. 23		
<p>이 연구는 인구주택총조사의 이용현황을 살펴보고, 이를 바탕으로 이용증진 방안을 제시하였다. 더불어, 현재 새로이 거론되는 센서스 방법들을 우리나라에 적용할 수 있는지를 살펴보고 적용방안 제안하였다. 이러한 방법에는 대표적으로 순환조사와 행정자료 기반의 센서스가 있다.</p> <p>먼저 인구주택총조사의 이용현황을 주 이용자인 공무원을 대상으로 설문조사를 통해 살펴본 결과, 지방자치단체 공무원이 중앙부처 공무원에 비해 이용하는 숫자도 많았고, 이용빈도도 높았다. 이들의 개선요구는 소지역 통계, 세부 통계, 시의성 있는 통계, 이용편의 등이 있었다. 이런 요구를 바탕으로 본 연구는 현재 공표 수준보다 작은 규모의 공표집계구 설정과 사용자 차별화한 원자료(microdata) 제공 서비스를 제안하였다.</p> <p>새로운 총조사 방안으로 행정자료 기반의 센서스와 순환조사를 살펴보았다. 먼저 행정자료 기반 센서스의 가능성에 대해 주민등록과 총조사 자료를 읍면동 수준에서 비교해 본 결과 인구수와 가구/세대수에서 큰 차이를 보인 지역이 상당히 많았다. 또한 건축물과 관련된 행정자료를 살펴본 결과 주소체계의 문제, 정보의 누락, 부정확한 정보 등으로 현재 상태에서는 그대로 총조사를 대체할 수 없었다. 이에 신축, 개축 등 건축물의 변동사항은 행정자료로 확인하여 변동이 없을 경우 통계청이 과거 2000년과 2005년에 조사한 가구와 주택부문 자료를 그대로 이용하는 방안을 제시하였다.</p> <p>순환조사의 경우, 현재는 프랑스식의 전수조사보다는 미국식의 표본조사가 더 실현 가능한 방안으로 보았다. 다만 표본조사를 5년 동안의 순환조사로 대체할 경우 2010년 총조사를 중심으로 2009년과 2011년에 시험조사가 있어야 하며, 그 이전에 표본설계와 함께 새로운 표집단위 개발이 필요하다. 또한 다른 조사들과 조정도 필요하다.</p> <p>마지막으로 총조사에서 행정자료의 활용은, 더 나아가 행정자료 기반의 센서스는 장기적으로 반드시 추구할 수 밖에 없는 통계생산방식임을 강조하였다.</p>			

Project Summary

Title of Project	A Methodological Study of the Population and Housing Census of 2010		
Key Words	Population and Housing Census, Survey Methods, Methodology		
Institute	Korean Association of Survey Research	Project Leader	LEE, Kun
Project Period	2007. 3. 27 ~ 2007. 11. 23		

This research has dual purposes: one is to assess the utility of the Population and Housing Census, and the other is to evaluate newly developed census methodology, namely register-based census and rolling census.

For the first task, we conducted on-line and off-line user survey and found out that local government officials use census statistics more than central government officials do. The other major finding is about their demands in the census, and they are small area statistics, detailed statistics, and easy access to microdata. We suggest the census needs to develop a new unit of aggregation that is smaller than the present size, and differentiated services of microdata for public use and confidential data.

For the second task, we examined the model of census items match with registers. We compared the aggregate data of the 2005 census and the population register at the end of 2005. The comparison entails many places with big differences in the number of population and households. We also examined several administrative registers related to buildings, and found out that they are incomplete, inaccurate, and hard to identify.

We also reviewed the French rolling census and American rolling survey, and suggested that to conduct rolling survey in the future, it is necessary to schedule several pilot surveys next year, and comparison study beginning from 2009. Secondly, it is important to develop a sampling unit for sample surveys with more consistent size. Thirdly, harmonizing with other sample surveys is not a small problem.

All these different methods in mind, however, we emphasize that utilizing administrative records in the short run, and pursuing a register based census in the long run are not options but the goals that we must achieve.

목차

제1장. 서론	1
제2장. 인구주택총조사 활용에 대한 조사	6
제1절. 서론	6
제2절. 조사방법	7
1. 조사대상	7
2. 표본	7
3. 조사방식	8
제3절. 응답자의 특성 및 분포	14
1. 응답률	14
2. 응답자의 특성	16
제4절. 인구주택총조사 자료 이용 현황	22
1. 일반적인 통계 자료 이용	22
2. 인구주택총조사 자료 이용	24
제5절. 인구주택총조사 자료 이용 만족도	35
1. 전반적 만족도	35
2. 항목별 평가	36
3. 소속기관별 평가	37
4. 자료 이용강도별 평가	38
제6절. 인구주택총조사에 대한 제안	39
1. 인구주택총조사에 대한 요구	39
2. 희망 추가조사항목	40
제7절. 소결	42
제3장. 인구주택총조사 활용 증진 방안	45
제1절. 서론	45
제2절. 외국의 통계 집계 및 공표구역	47
1. 전반적인 경향과 특징	47

2. 미국의 통계구역 설정 현황	50
3. 영국의 통계구역 설정 현황	53
4. 캐나다의 통계구역 설정 현황	59
5. 호주의 통계구역 설정 현황	63
6. 일본의 통계구역 설정 현황	67
제3절. 우리나라의 통계구역 및 공표구역 지정 현황	70
1. 통계권역체계	70
2. 각 통계권역의 개념 및 설정방법	72
제4절. 원자료 활용 증진 방안	78
1. 원자료 수요 증가와 자료 서비스 문제	78
2. 원자료 서비스의 전반적 경향	80
3. Public Use Microdata	82
4. Confidential data	84
제5절. 소결	88
제4장. 행정자료 기반의 센서스	91
제1절. 서론	91
제2절. 외국의 센서스	92
1. 다양한 센서스 방식	92
2. 행정자료의 보완적 활용	96
3. 행정자료 기반의 센서스	100
제3절. 우리나라 행정자료 기반 센서스의 가능성	118
1. 이론적 모형	118
2. 인구주택총조사에서 행정자료의 활용가능성	120
제5장. 인구주택총조사자료와 주민등록자료 비교	126
제1절. 서론	126
제2절. 인구 비교	127
1. 연구배경	127
2. 자료	128
3. 비교 결과	129

4. 소결 및 논의	142
제3절. 가구와 세대 비교	143
1. 가구와 세대의 개념 비교	143
2. 비교 결과	145
3. 소결	160
제4절. 연령별 비교	162
1. 서론	162
2. 비교 결과	163
3. 소결	186
제5절. 개인수준의 매칭	187
1. 자료	187
2. 매칭 방법	194
3. 매칭 결과	197
4. 소결 및 시사점	210
제6장. 인구주택총조사에서 건축물대장의 활용 가능성	213
제1절. 서론	213
제2절. 건축물 관련 행정자료	214
1. 건축물대장	214
2. 기타 건축물 관련 행정자료	222
제3절. 건축물대장 활용의 문제	226
1. 건축물 대장의 문제점	226
2. 건축물대장과 관련 자료 연계의 문제	234
제4절. 행정자료 간 매칭 시 문제	244
1. 행정자료 간 연계의 문제	244
2. 행정자료의 정확도 문제	245
3. 동일항목 비교 시 문제	246
4. 건축물 정보에 대한 총체적 현황 파악의 어려움	248
제5절. 시사점	249
1. 건축물 정보 간 연계의 문제	249

2. 개별 자료들의 정확도 문제	250
3. 동일항목에 대해서도 비교 불가능한 문제	251
4. 건축물에 대한 총체적 현황 파악의 어려움	251
제7장. 순환총조사	252
제1절. 개요	252
제2절. 외국의 사례	254
1. 프랑스 순환총조사	254
2. 미국지역사회조사	260
제3절. 순환총조사 도입 시 검토사항	267
1. 총조사 결과물	268
2. 사용자 수용성	268
3. 연속시행	269
4. 총조사의 브랜드 가치 하락	269
5. 포함 범위	270
6. 법령 제정	270
7. 비용	270
제4절. 우리나라 실정에 맞는 순환총조사 방법 검토	271
1. 예산 및 인력	271
2. 순환총조사 방법론 개발	271
3. 지방자치단체의 협조 필요	273
4. 순환총조사에 대한 홍보 및 법령 제정	273
제5절. 소결	274
제8장. 결론 및 제언	275
제1절. 인구주택총조사의 활용성	275
1. 설문조사	275
2. 활용 증진 방안	276
제2절. 행정자료 이용	278
제3절. 순환(총)조사	282
제4절. 제언	285

제9장. 참고문헌	289
<부록 1> 온라인 설문지	299
<부록 2> 유럽 국가들의 센서스 방식	305
<부록 3> 노르웨이 2001년 센서스 행정자료 활용	306
<부록 4> 세대의 법적 정의	308
<부록 5> 주민등록에 관한 법규	310
<부록 6> 지방자치단체에 관한 규정	311
<부록 7> 인구주택총조사에서 GIS 활용 방안	312
제1장. 조사구 설정과 GIS	312
제1절. 전통적 방식의 조사구 설정업무	312
1. 개요	312
2. 거쳐 파악과 조사구 설정	313
3. 준비조사와 조사구 확정	316
제2절. GIS 기반의 조사구 설정자료 관리 및 활용	318
1. GIS를 기반의 조사구 설정 및 자료의 DB 구축	318
2. GIS 기반으로 다른 업무와 연계 및 결과 활용	321
제2장. 기초단위구 개요 및 활용	323
제1절. 기초단위구의 개념 및 확정 기준	323
제2절. 기초단위구 지정 및 관리 현황	324
제3절. 기초단위구 번호부여 체계	326
제4절. 기초단위구의 활용	328
1. 도시화지역 설정	328
2. 도시권설정	329
3. 기타	330
제3장. 모바일 GIS 개요 및 활용사례	331
1. 모바일 GIS를 이용한 현장업무의 효율성 확보	331
2. 모바일 GIS의 개요	332
3. 모바일 GIS 적용 사례	334
<부록 8> 2010년 인구주택총조사 추진일정	339

표 차례

<표 2-1> 중앙부처별 추출 인원 및 확보 이메일 수	9
<표 2-2> 지방자치단체별 추출 인원 및 확보 이메일 수	10
<표 2-3> 조사집단별 이메일 확보비율	10
<표 2-4> 응답대상에 따른 온·오프라인 응답률 분포	15
<표 2-5> 이번조사와 2003년 조사의 비교	16
<표 2-6> 응답자의 성별 분포	16
<표 2-7> 응답자의 연령 분포	17
<표 2-8> 조사방법별 응답자의 연령 분포	17
<표 2-9> 중앙부처별 발송건수 및 응답률	19
<표 2-10> 지방자치단체별 발송건수 및 응답률	21
<표 2-11> 응답자의 소속기관별 직급 분포	22
<표 2-12> 소속기관별 통계 자료 이용 빈도	23
<표 2-13> 조사방법별 통계 자료 이용 빈도	23
<표 2-14> 소속기관별 인구주택총조사 자료 이용경험	24
<표 2-15> 조사방법별 인구주택총조사 자료 이용경험	25
<표 2-16> 인구주택총조사 자료별 이용경험률 (복수응답)	26
<표 2-17> 시도별 추가표본조사항목 및 표본수	27
<표 2-18> 인구주택총조사 이용 자료형태 (복수응답)	29
<표 2-19> 조사방법별 인구주택총조사 이용 자료형태 (복수응답)	30
<표 2-20> 인구주택총조사 자료의 주된 이용용도 (복수응답)	31
<표 2-21> 인구주택총조사 자료이용 빈도	32
<표 2-22> 조사방법별 인구주택총조사 자료 이용빈도	33
<표 2-23> 인구주택총조사 자료활용의 업무 영향력 평가	34
<표 2-24> 자료이용강도별 평가	34
<표 2-25> 인구주택총조사 자료에 대한 전반적 만족도	36
<표 2-26> 인구주택총조사 자료 항목별 평가	37

<표 2-27> 소속기관별 인구주택총조사 자료 평가	38
<표 2-28> 인구주택총조사 자료 이용강도별 평가	39
<표 2-29> 인구주택총조사 개선 요청사항	41
<표 2-30> 인구주택총조사 희망 추가 조사항목	42
<표 3-1> 외국의 통계 관련 구역의 특징*	49
<표 3-2> 지역별 블록그룹의 인구 기준	52
<표 3-3> 미국의 RDC 운영 현황	85
<표 4-1> 센서스 방식에 따른 국가 분류	93
<표 4-2> 노르웨이의 항목별 행정자료 대체 연도	102
<표 4-3> 스웨덴의 항목별 행정자료 대체 연도	105
<표 4-4> 덴마크의 중앙인구등록번호 체계	107
<표 4-5 > 덴마크의 주소번호 체계	107
<표 4-6> 덴마크의 센서스 관련 주요 사건 연혁	108
<표 4-7> 덴마크의 항목별 행정자료 대체 연도	110
<표 4-8> 덴마크 통계관련 법	111
<표 4-9> 핀란드의 중앙인구등록번호 체계	112
<표 4-10> 핀란드의 건물/주거 번호 체계	113
<표 4-11> 핀란드의 센서스 관련 주요 사건 연혁	113
<표 4-12> 핀란드의 항목별 행정자료 대체 연도	116
<표 4-13> 핀란드의 정보보호 관련 법	117
<표 4-14> 2005년 센서스 항목의 행정자료 활용가능성	121
<표 5-1> 주민등록인구와 총조사인구의 시도별 차이, 1995, 2000, 2005년	132
<표 5-2> 주민등록과 인구주택총조사 인구의 차이가 큰 시군구	135
<표 5-3> 읍면동 수준의 총조사인구/주민등록인구 분포	137
<표 5-4> 주민등록대 인구주택총조사 인구의 차이가 30% 이상인 읍면동 ...	140
<표 5-5> 세대와 가구 개념 비교	146
<표 5-6> 주민등록세대와 총조사가구의 시도별 차이	148
<표 5-7> 총조사가구가 주민등록세대보다 20% 이상 적은 경우-총가구	152
<표 5-8> 총조사가구가 주민등록세대보다 20% 이상 적은 경우-일반가구 ...	153

<표 5-9> 읍면동 수준의 총조사가구/주민등록세대 분포	157
<표 5-10> 주민등록세대와 총조사가구의 읍면동별 차이율-총가구	158
<표 5-11> 주민등록세대와 총조사가구의 읍면동별 차이율-일반가구	159
<표 5-12> 연령대별 주민등록인구와 총조사인구의 차이와 차이율	164
<표 5-13> 차이율 및 연령대 절대차이율 평균의 시도별 비교	169
<표 5-14> 연령대 절대차이율 평균 기준 시군구별 차이율 : 상위 10개 ..	175
<표 5-15> 20-24세 연령대 차이율 기준 시군구별 차이율	176
<표 5-16> 연령대 절대차이율 평균 기준 읍면동 차이율 : 상위 10개	181
<표 5-17> 세대주관계	189
<표 5-18> 시범조사 자료에 입력된 값	190
<표 5-19> 가구주와의 관계	191
<표 5-20> 시도(시군구)별 총인구	191
<표 5-21> 성별 인구	194
<표 5-22> 방법1을 이용한 매칭 결과	197
<표 5-23> 방법1을 이용한 성별 매칭 결과	198
<표 5-24> 방법1을 이용한 연령별 매칭 결과	199
<표 5-25> 방법1을 이용한 연령별 매칭 일치율	200
<표 5-26> 방법2를 이용한 매칭 결과	201
<표 5-27> 방법2를 이용한 성별 매칭 결과	203
<표 5-28> 방법2를 이용한 연령별 매칭 결과	203
<표 5-29> 방법2를 이용한 연령별 매칭결과	204
<표 5-30> 방법1을 이용한 매칭결과	205
<표 5-31> 방법1을 이용한 매칭결과	207
<표 5-32> 방법2를 이용한 부산광역시 해운대구 우동의 매칭결과	208
<표 5-33> 방법2를 이용한 매칭결과	209
<표 5-34> 부산광역시 해운대구 전체의 매칭결과	210
<표 6-1> 건축물대장의 종류	215
<표 6-2> 건축허가·신고 대상	218
<표 6-3> 건축물 대장별 주요 항목	219

<표 6-4> 위법 건축물 관리 대장별 항목	224
<표 6-5> 등기부등본 정보	225
<표 6-6> 건축물대장의 주요 항목별 누락률-92년도이후 사용승인분	227
<표 6-7> 건축물대장의 주요 항목별 누락률-층별자료, 총괄표제부	228
<표 6-8> 건축물 도형과 건축물대장의 연계	236
<표 6-9> 건축물도형과 건축물대장의 연계율	238
<표 6-10> 건물 도형과 건축물대장의 연계	239
<표 6-11> 건축물 관련 자료의 구조분류체계 비교	246
<표 6-12> 건축물관련 자료의 용도분류체계 비교	247
<표 6-13> 서울시 대장별 건축물 수 비교	249
<표 7-1> 프랑스 자치구 현황	256
<표 7-2> 무응답 비율	260
<표 7-3> 표본 추출률	262
<표 7-4> 일년 추정치와 다년 추정치 : 증가 추세 항목	265
<표 7-5> 일년 추정치와 다년 추정치 : 등락이 있는 항목	265
<표 부록7-1> 인구주택총조사의 거처 구분	314
<표 부록7-2> 조사구 분류	315
<표 부록7-3> 공동주택 기본정보 수집표(양식) 예시	318
<표 부록7-4> 가구명부 예시	318
<표 부록7-5> 대구역과 기초단위구(소구역) 구획 기준	324
<표 부록7-6> 시도별 기초단위구 현황 및 변동내역	325
<표 부록7-7> 기초단위구 특성번호 분류 기준	327
<표 부록7-8> 일부지역의 기초단위구별 가구수 (2000년 기준)	328
<표 부록7-9> 모바일 GIS 응용구분	334
<표 부록7-10> 위치기반서비스(LBS)와 차량항법시스템(ITS)	336
<표 부록7-11> 해외 모바일시스템 활용 현황	337

그림 차례

(그림 2-1) 조사과정 요약	11
(그림 2-2) 온라인조사 안내문	12
(그림 2-3) 웹 서버이 초기화면	13
(그림 2-4) 웹 서버이 설문화면	13
(그림 2-5) 중앙부처별 응답자수	18
(그림 2-6) 지방자치단체별 응답자수	20
(그림 2-7) 인구주택총조사 자료별 이용경험률 (복수응답)	26
(그림 2-8) 인구주택총조사 자료이용 형태(복수응답)	28
(그림 3-1) 미국의 통계구역 구분 체계	51
(그림 3-2) 미국의 소지역 통계구역의 구분과 번호체계	53
(그림 3-3) 영국의 통계구역 구분 체계	54
(그림 3-4) 조사구역과 공표구역의 분리	57
(그림 3-5) 캐나다 통계구역(Statistical Area) 체계	61
(그림 3-6) 캐나다의 블록 설정	63
(그림 3-7) 호주 통계구역(Statistical Area) 체계	65
(그림 3-8) 우리나라의 통계권역 체계	71
(그림 3-9) 통계지역체계	72
(그림 3-10) 조사구 예시 도면	73
(그림 3-11) 사례지역별 집계구 설정결과	76
(그림 3-12) 도시화지역 확정 사례-용인시	77
(그림 3-13) Microdata 이용자 증가 현황	79
(그림 3-14) IPUMS의 원자료 예시	83
(그림 3-15) IPUMS 참여국가	83
(그림 4-1) 스웨덴 행정자료 기반 센서스 구성도	104
(그림 4-2) 핀란드 행정등록 기반 센서스 구성도	115
(그림 4-3) 행정자료 연계를 위한 행정자료의 기본틀	120

(그림 4-4) 주민등록과 인구주택총조사 항목의 대응 예시	122
(그림 4-5) 건축물대장과 인구주택총조사 항목의 대응 예시	123
(그림 5-1) 시도별 주민등록인구 대비 인구주택 총조사인구의 비율	131
(그림 5-2) 시군구 수준의 주민등록인구와 총조사인구 차이율 분포	133
(그림 5-3) 주민등록과 총조사 인구의 시군구별 차이율 상자도표	134
(그림 5-4) 주민등록 세대와 총조사 가구의 읍면동별 차이율 분포	136
(그림 5-5) 주민등록 세대와 총조사 가구의 시군구별 차이율 상자도표	138
(그림 5-6) 서울의 주민등록 대비 인구주택총조사 인구비율의 분포	142
(그림 5-7) 주민등록세대와 총조사가구의 시군구별 차이율-총가구	151
(그림 5-8) 주민등록세대와 총조사가구의 시군구별 차이율-일반가구	151
(그림 5-9) 주민등록세대와 총조사가구의 시군구별 차이율 상자도표-총가구	154
(그림 5-10) 주민등록세대와 총조사가구의 읍면동별 차이율-총가구	155
(그림 5-11) 주민등록세대와 총조사가구의 읍면동별 차이율-일반가구	156
(그림 5-12) 주민등록세대와 총조사가구의 읍면동별 차이율 상자도표-총가구	160
(그림 5-13) 연령대별 차이와 차이율	165
(그림 5-14) 지역단위별 연령대별 차이율 평균 비교	166
(그림 5-15) 지역단위별 연령대별 차이율 표준편차 비교	167
(그림 5-16) 연령대별 차이율 : 광역시 비교	170
(그림 5-17) 연령대별 차이율 : 도 비교	170
(그림 5-18) 15-19세 연령대 차이율	171
(그림 5-19) 20-24세 연령대 차이율	171
(그림 5-20) 25-29세 연령대 차이율	172
(그림 5-21) 30-34세 연령대 차이율	172
(그림 5-22) 35-39세 연령대 차이율	173
(그림 5-23) 70-74세 연령대 차이율	174
(그림 5-24) 절대차이율 평균 기준 상위 10개 시군구별 연령대별 차이율	175
(그림 5-25) 15-19세 연령대 시군구별 차이율	177
(그림 5-26) 20-24세 연령대 시군구별 차이율	178
(그림 5-27) 25-29세 연령대 시군구별 차이율	178

(그림 5-28) 30-34세 연령대시군구별 차이율	179
(그림 5-29) 35-39세 연령대 시군구별 차이율	179
(그림 5-30) 70세 이상 연령대 시군구별 차이율	180
(그림 5-31) 절대차이율 평균 기준 상위 10개 읍면동 연령대별 차이율	182
(그림 5-32) 20-24세 읍면동별 차이율	183
(그림 5-33) 15-19세 읍면동별 차이율	183
(그림 5-34) 25-29세 읍면동별 차이율	184
(그림 5-35) 30-34세 읍면동별 차이율	184
(그림 5-36) 35-39세 읍면동별 차이율	185
(그림 5-37) 70세 이상 읍면동별 차이율	185
(그림 5-38) 주민등록자료의 형식	188
(그림 5-39) 시범조사자료의 일부	189
(그림 5-40) 충청북도 보은군의 연도별 인구 변화	192
(그림 5-41) 충청북도 보은군의 인구	193
(그림 5-42) 방법1을 이용한 매칭의 일치율 비교	198
(그림 5-43) 방법1을 이용한 매칭의 일치율 비교	200
(그림 5-44) 방법2를 이용한 매칭의 일치수 비교	202
(그림 5-45) 방법2 이용한 매칭의 일치율 비교	202
(그림 5-46) 방법2를 이용한 매칭의 일치율 비교	204
(그림 5-47) 방법1을 이용한 매칭의 일치수 비교	206
(그림 5-48) 방법1을 매칭의 일치율 비교	206
(그림 5-49) 방법2를 이용한 매칭의 일치수 비교	208
(그림 5-50) 방법1을 이용한 매칭의 일치율 비교	208
(그림 6-1) 건축물대장 관련 도면 예시-평면도	215
(그림 6-2) 건축물대장 관련 도면 예시-배치도	216
(그림 6-3) 건축물대장 작성 절차	217
(그림 6-4) 건축물 대장의 전산화된 형태 예시	220
(그림 6-5) 주건축물수와 건축물수 불일치	229
(그림 6-6) 주용도 분류가 다른 경우	229

(그림 6-7) 면적/건폐율/용적률 등이 다르거나 값이 없는 경우	229
(그림 6-8) 허가일자/착공일자/사용승인일자가 다른 경우	230
(그림 6-9) 필지번호와 건물의 관계 유형	235
(그림 6-10) 2차 연계 작업	238
(그림 6-11) 건축물도형과 대장간 1차 연계 결과 (창신동)	239
(그림 6-12) 건축물도형과 대장간 추가 연계 결과 (창신동)	240
(그림 6-13) 건축물도형과 대장간의 1차 연계 결과 (역삼동)	241
(그림 6-14) 건축물도형과 대장간의 추가 연계 결과 (역삼동)	242
(그림 6-15) 건축물도형과 대장간의 1차 연계 결과 (창천동)	243
(그림 6-16) 건축물도형과 대장간의 추가 연계 결과 (창천동)	244
(그림 부록3-1) 2001년 센서스 이전과 이후의 등록체계	306
(그림 부록3-2) 2001년 센서스의 자료 출처	307
(그림 부록7-1) 인구주택총조사 조사부분의 흐름	313
(그림 부록7-2) 조사구지(요)도 사용 방법 예시 (조사 전 → 조사 후)	317
(그림 부록7-3) 통계청의 기초단위구 관리 업무현황	320
(그림 부록7-4) 통계청의 조사구 관리 업무현황	320
(그림 부록7-5) 우리나라 기초단위구 일련번호 구성 체계	326
(그림 부록7-6) 도시화지역 획정 사례-용인시	329
(그림 부록7-7) 통계지리정보팀 향후 업무계획	330
(그림 부록7-8) 온라인 모바일 GIS 구조	333
(그림 부록7-9) 오프라인 모바일 GIS 구조	333

제1장. 서론

우리나라의 인구주택총조사는 1925년에 처음 시작되어 그 이후 매 5년마다 실시되어왔다. 지난 80년 동안 인구주택 총조사는 국가의 운영에 반드시 필요한 통계를 생산하는 조사로 인식되어 그다지 큰 어려움이 없이 시행되었다. 그러나 도시화에 따른 가구 형태의 변화 및 다양화로 전통적인 조사방법으로 파악하기 어려운 가구들이 증가하여 이를 보완할 조사방법이 필요해지고 있다. 또한 최근 개인정보에 대한 인식이 향상되면서 총조사에 대한 거부감이 조금씩 나타나기 시작하여 전통적인 조사방법을 보완할 수단이 필요해지고 있다.

이렇듯 사회 환경의 변화에 따라 조사의 어려움은 커지면서 조사비용도 증가하기 시작하였다. 또한 인구수의 증가와 핵가족화 및 가구 분할로 파생하는 가구수의 증가로 조사할 대상이 많아지는 것도 조사비용을 높이는 데 기여하였다. 이와 더불어 물가와 인건비 상승에 따른 조사단가상승도 조사비용을 높이고 있다. 가구당 비용은 1990년에 2,146원, 1995년에 4,597원, 2000년에 6,733원, 2005년에 8,875원으로 15년 사이에 4배 이상이 되었다(국가예산정책처, 2006). 총조사의 총비용도 1990년 216억 원에서, 1995년 542억 원, 2000년 834억 원, 2005년에는 1,290억 원으로 6배가량 급격하게 증가하였다(국가예산정책처, 2006). 조사비용의 증가는 비용효율적인 조사방법을 모색하는 계기가 되고 있으며, 이와 더불어 조사결과의 활용을 높이는 방안을 모색하는 계기도 되고 있다.

한편 통계수요의 변화도 총조사의 조사방법을 살펴보아야 할 이유다. 특히 지방자치단체가 행정의 독립성을 가지면서 지역통계에 대한 수요가 증가하였다. 이에 2005년 총조사에서 16개 시도별로 표본조사항목을 마련하였다. 그 결과 조사 관리의 어려움도 있었으며 조사비용도 증가하여 이에 대한 보완이 필요해졌다(통계청, 2006a)

새로운 인구센서스 방법의 탐색은 단지 우리나라에서만 일어나는 현상이 아니다. 오랜 동안 정기적으로 인구센서스를 해오던 유럽과 미국 등 선진국에서도 조사비용 증가, 조사공표시기의 시의성, 응답자 부담과 응답거부의

증가 등으로 새로운 센서스 방식을 모색하고 있다(Jensen, 2000, Laihonen, 2000). 특히 1980년 이후 전통적 방식으로 인구센서스 통계를 생산하지 못한 네덜란드나 독일의 경험은 많은 국가들이 새로운 센서스 방식을 찾는 계기가 되었다. 1970년대에 들어 네덜란드 국민들 사이에서 응답거부에 대한 우려가 커지면서 1971년 마지막 인구센서스 이후 지금까지 전수조사를 시행하지 못하였다(van der Laan, 2000). 독일 또한 1980년대에 센서스에 대한 거부감이 증가하여 1987년에 전통적 방식으로 센서스를 한 이후 아직 센서스를 실시하지 못하고 있다(Bierau, 2000; Eppmann, Krügener, and Schäfer, 2006). 이들보다는 좀 덜 심각하지만 1990년대 미국은 센서스 통계의 시의성을 높이고자 다양한 방법들을 검토한 다음, 기존 센서스의 표본조사를 순환조사(rolling survey)로 바꾸는 방안을 제안했다(Edmonston and Schultze, 1995). 비슷한 시기에 영국은 센서스의 정확성을 높이기 위해 표본조사를 통해 전수조사의 인구수를 보정하는 One Number Census를 개발하였다(ONS, 2000, 2001).

새롭게 제기되는 센서스 방식으로는 크게 순환조사(rolling survey), 행정자료를 기반으로 하는 센서스(register based census), 표본조사와 행정자료를 혼합하는 방식이 있다(ONS, 2003). 또한 전통적으로는 모든 국민의 머릿수를 하나씩 세는 방식에만 의존하였으나, 최근에는 통계적 추정기법을 다양하게 활용하는 방식으로 변화하고 있는 추세다. 먼저 순환조사는 현재 프랑스와 미국에서 준비하고 있는 방식으로 전체 조사대상을 매년 나누어 조사하여 센서스 연도에 전체 자료를 확보하는 방법이다. 프랑스는 전수조사를 순환조사로 대체하여 이를 순환총조사(rolling census)라고 부른다. 한편 미국은 2010라운드 센서스에서는 전수조사와 함께 시행하던 표본조사를 순환조사인 American Community Survey로 대체할 예정이다. 네덜란드와 독일에서는 표본조사와 행정자료를 적절히 혼합하는 방식을 시험하고 있다. 행정자료 기반의 센서스는 주로 북구 유럽에서 발달하였으며, 그 가운데도 덴마크와 핀란드가 가장 잘 알려져 있다. 그 결과 유럽에서는 전통적인 조사방법을 사용하는 나라가 2000라운드 센서스에서는 80%였으나 2010라운드 센서스에서는 55%로 감소할 전망이다(이지연, 2007a).

새로운 센서스 방식들 가운데 가장 주목받는 것이 행정자료 기반의 센서스다. 최근 발달한 ICT(Information Communication Technology) 기술로 행정자료의 축적 및 관리가 수월해지고, 이들을 연계하여 통계를 생산할 수 있는 기반이 마련되었기 때문이다. 행정자료 기반 센서스의 장점은 조사를 하지 않기 때문에 막대한 조사비용이 절감되고, 응답자의 부담도 없으며, 응답 거부 또한 없다. 더욱이 행정자료를 기반으로 하면 5년이나 10년의 센서스 주기와 관계없이 필요한 때에 자료를 모아 통계를 생산할 수 있다는 장점이 있다.

IT 강국으로 알려진 우리나라에서도 행정자료 기반의 센서스에 대한 관심이 높다. 우리나라는 행정자료 기반의 국가통계 작성과 관련하여 몇 가지 중요한 기반을 갖추고 있기 때문이다. 먼저, 1962년부터 시행한 주민등록제도로 40년 이상 시행되어 이제는 거의 오류가 없으며, 모든 국민 개개인을 파악할 수 있는 주민등록번호로 모든 행정 및 공식 기관에서 사용하여 자료 사이의 연계가 수월하다. 이와 더불어, 우리나라의 높은 정보화 수준과 2000년부터 시작된 전자정부사업으로 행정등록 자료를 연계·통합할 수 있는 기반이 잘 갖추어져 있다. 더욱이, 모든 행정에서 주민등록을 사용하고 있기 때문에 이로 센서스를 대체할 수 있다는 견해가 점점 확산되고 있다.

그렇지만 지금 당장 주민등록으로 인구센서스를 대체할 수 있는 것은 아니다. 먼저, 인구센서스와 주민등록의 인구 및 가구는 개념이나 파악하는 방식이 똑같지 않다. 따라서 주민등록자료로 인구센서스를 대체하려면 인구센서스에 맞도록 주민등록의 정보를 조정해야 한다. 또한 지금까지 축적해온 인구센서스 자료와 시계열을 맞추기 위한 조정도 필요하다. 이러한 조정에는 많은 시간이 소요된다. 이것은 두 자료의 비교·대조가 그다지 쉽지 않을뿐더러, 인구센서스가 5년에 한번 있는 조사이기 때문이다. 조정을 위한 조사 설계가 인구주택총조사에 포함되어야 하며, 그 조사결과를 바탕으로 조정 작업을 진행해야 한다. 따라서 2번의 인구센서스를 거친다면 적어도 10년이 소요된다. 이미 행정자료 기반으로 인구센서스를 집계하는 나라에서 새 방식의 도입에 15 - 20년 이상의 기간이 소요된 것을 보면 그러한 조정 작업이 얼마나 지난한 일인지를 알 수 있다.

우리는 아직 인구주택총조사를 하기 어려운 아주 긴박한 상황에 처한 것은 아니다. 그렇지만 외국의 경우를 보면 그러한 상황은 점점 우리에게 다가오고 있는 것으로 보인다. 이와 함께 국민들과 행정부의 통계에 대한 요구도 달라지고 있다. 이런 문제를 예견하고 앞으로 나타날 문제에 대처하기 위해 이 연구는 앞으로 우리가 인구주택총조사를 어떤 방식으로 수행해야 할지를 살펴보고자 한다. 그렇다고 이 연구가 미래의 인구주택총조사의 방향을 결정하는 것은 아니다. 다만, 세계의 흐름, 우리의 상황 등에 비추어 인구주택총조사가 가야할 큰 방향을 제안하는 것이다.

이런 목적을 위해, 이 연구는 먼저 현재 인구주택총조사가 어떻게 사용되는지를 설문조사를 통해 살펴보았다. 설문조사에서는 중앙부처와 지방자치단체의 통계관련 공무원과 학계 및 연구기관의 연구원을 대상으로 인터넷 조사와 우편조사를 동시에 진행했다. 다음은 설문조사 결과를 토대로 우리나라의 인구주택총조사가 좀 더 활용되기 위해서는 어떤 것들이 필요한지를 제안하였다. 여기서는 특히 소규모 지역으로 통계를 공표할 수 있는 통계집계 및 공표구역과 원자료 이용을 중점적으로 다루었다. 특히 통계구역에서 중요한 구현도구가 Geographical Information System(GIS)이다. 이의 유용성을 좀 더 부각하기 위해 부록에 통계 GIS 관련 자료를 담았다.

새로운 센서스 방식에서는 먼저 행정자료 기반의 센서스에 대한 외국의 사례들을 살펴본 다음, 이론적 모형을 제시하고, 김설희·강계화(2005)가 제안한 행정자료로 대체가 가능한 총조사 항목에 관한 가설을 검증하고자 하였다. 이를 위해 주민등록 집계자료와 총조사 집계자료를 읍면동 수준에서 비교하였다. 분석에서는 읍면동 지역별로 인구수, 세대와 가구수, 연령별 인구수를 비교하였다. 이와 더하여, 개인수준에서 총조사 시범자료와 같은 시기의 주민등록 자료를 매칭하고자 하였다. 또한 건축물 관련 행정자료의 현황을 살펴보고, 그들의 연계 가능성, 자료의 정확도 등을 면밀히 들여다 본 다음, 앞으로 이 자료를 통계적으로 이용하기 위해서는 어떤 문제가 있는지 어떤 작업이 필요한지를 찾아보았다.

순환조사는 아직 완료된 경우가 없고 프랑스와 미국에서 진행 중인 상태에서 두 나라를 중점적으로 살펴보았다. 이를 토대로 순환조사를 할 경우 어

면 것들을 갖추어야 하는지를 제시하고, 우리나라에 적합한 방식이 무엇인지를 찾아보고자 하였다.

결론에서는 본 연구를 간략하게 요약한 다음, 행정자료 기반 센서스와 순환조사에 대해 통계적 논의와 함께 시대 상황적 논의를 전개하였다. 특히 두 가지 방법에서 현재 우리가 주목해야 할 것들이 무엇이며, 더 나아가 미래의 한국에서 통계청이 택해야 할 주된 방법이 무엇인지를 제시하고자 하였다.

마지막으로, 이 연구는 21세기 인구주택총조사의 방향을 정하는 데 필요한 탐색작업이다. 따라서 일반 학술연구에 비해 연구 범위가 방대하고, 연구 방법이 다양하다. 이런 덕분에, 이 연구는 일반 학술연구가 갖는 깔끔하고 논리 정연한 모습을 갖기가 어렵다. 반면, 넓은 범위에서 다양한 문제들을 다양한 방식으로 다루어 많은 정보를 갖고 있다. 이런 측면에서 이 연구에서는 본래 이 연구가 추구했던 탐색적 연구의 성격이 잘 드러나고 있다고 할 수 있다. 탐색의 목적은 물론 새로운 센서스 방식을 찾아가는 것이며, 이 연구는 그 첫 걸음이다.

제2장. 인구주택총조사 활용에 대한 조사

제1절. 서론

인구주택총조사는 우리나라의 인구, 주택, 가구수를 파악하여 가장 단순하면서도 가장 중요한 통계를 생산한다. 이들은 국가의 정책수립이나 집행에 필요할뿐더러, 민간기업의 제품 생산이나 판매 계획 수립에 유용한 통계를 제공한다. 또한 인구주택총조사의 전수조사결과는 각종 국가통계를 생산하기 위한 표본의 모집단으로 활용되고, 표본조사결과는 그러한 표본을 추출하는데 이용되고 있다. 이외에도 표본조사결과는 경제정책, 주택정책, 교통정책 등에 활용되어왔다.

이렇듯 다양한 용도에도 불구하고, 인구주택총조사 자료가 통계청 이외의 기관에서 반드시 필요한 자료로 인식되는 경우는 그다지 많지 않다. 국가의 행정에서 인구주택총조사 자료를 직접 사용하는 경우가 그다지 많지 않기 때문이다. 특히 인구주택총조사의 대표 격인 인구수를 행정에서는 주민등록 인구로 사용하기 때문에 더욱 그러하다.

이런 배경에서 인구주택총조사 자료를 어디에 활용하는지 알아 볼 필요가 발생하였다. 더불어 활용도를 높이기 위한 방안도 마련할 필요가 생겼다. 이 연구는 그러한 현황과 활용 방안을 마련하기 위해 통계청의 2005년 인구주택총조사가 활용에 관한 설문조사를 실시하였다.

설문조사에서는 중앙부처와 지방자치단체, 연구기관을 대상으로 총조사 자료가 어떤 용도로 사용되는지, 어떤 통계가 유용한지, 조사 및 공표되는 시기는 적절한지 등을 파악하였다. 이와 함께, 2005년 총조사의 표본조사에서는 지방자치단체가 제안한 항목들도 조사했는데, 이의 활용성에 대해서도 파악하였다. 특히 온라인 조사도 함께 실시하여, 지방자치단체별 질문에 대해 혼란 없이 응답할 수 있도록 하였다. 또한 인구주택총조사에 대한 개선 요구나 필요한 사항들을 개방형 질문으로 알아보았다.

제2절. 조사방법

1. 조사대상

인구주택총조사의 본래 취지가 국가의 정책수립이나 집행을 위한 기초자료 제공이다. 이런 취지에 맞춰 본 조사는 중앙부처 및 지방자치단체 공무원, 중앙부처와 지자체 산하 연구기관 연구원, 대학교수, 전문조사기관 연구원 등 연구직으로 조사대상을 한정하였다. 중앙부처와 지방자치단체 공무원들 가운데 3급보다 높은 직급의 사람들은 주로 조직 및 업무를 관리하는 역할을 하여 실제로 정책수립과 집행에 필요한 자료를 사용하는 일이 드물다. 이런 이유에서 조사대상인 공무원은 3급 이하의 직급에 있는 사람들로 하였다. 공무원 이외에 연구기관이나 학자들 가운데서 연구를 위해 인구주택총조사의 자료를 이용하는 경우가 있어 이들도 조사대상에 포함하였다.

2. 표본

이번 조사의 대상자 대부분이 중앙부처 및 지방자치단체 소속 공무원들이 차지하는 만큼, 표본이 몇 개의 부처나 지자체로 편중될 경우, 조사결과의 심각한 왜곡(bias)을 초래할 우려가 있다. 이를 방지하기 위해 모든 부처와 지방자치단체에서 가급적 많은 리스트를 확보하여 표본을 추출하였다. 표집틀(sampling frame)은 각 중앙부처 및 지방자치단체별 전화번호 명단이나 홈페이지에서 각 부서의 팀장 및 그 이하 팀원의 명단을 사용하였다. 조사대상 중앙부처 공무원은 인구주택총조사의 담당 주무부처인 통계청을 제외한 정부의 39개 부처 및 위원회에서 각각 30명~40명씩 1,240명을 무작위로 추출하였다. 명단사용이 극히 제한적이었던 기획예산처, 식품의약품안전청, 철도청을 제외한 부처 및 위원회에서 최소 표본수 30명을 확보하였다. 부족한 표본을 보완하기 위해 홈페이지에서 상대적으로 명단 사용이 자유로웠던 건설교통부, 교육인적자원부 같은 부처들에서는 표본을 40명씩을 추출하였다. 이렇게 선정한 총 표본 수는 1,240명으로, 표본을 추출한 부처의 명단과 각

부처별로 추출한 표본 수를 아래 <표 2-1>에 제시하였다.

지방자치단체에 소속된 공무원들도 중앙부처와 같은 방식으로 표본추출을 하였다. 지방자치단체에서 추출한 총 표본 수는 539명으로, 표본을 추출한 지자체의 명단과 각 단체별로 추출한 표본 수를 아래 <표 2-2>에 제시하였다.

대학교수나 연구기관 등의 경우, 분야별로 학회의 회원 명단과 전국 주요 대학의 교수 명단이 표집틀이 되기는 하지만 이들의 대부분이 통계자료를 활용하는 것은 아니다. 이런 이유에서 이들 전체를 대상으로 하는 것은 심각한 응답률 저하를 가져오고 조사비용을 크게 높이는 문제를 일으킬 수 있다. 이런 문제를 해결하기 위해 이 연구는 지난 5년 동안 통계청에서 인구주택 총조사 자료를 사용하였던 연구자들의 명단을 표본으로 하였다. 이렇게 확보된 연구관련 종사자들은 90명 이었다.

3. 조사방식

조사는 크게 이메일을 이용한 온라인조사와 우편을 이용한 오프라인 조사로 나누어 진행하였다. 이렇게 두 가지 조사를 병행한 이유는 첫째는, 응답률을 높이기 위한 방안으로, 둘째는 이메일 주소를 알 수 없는 경우 때문이었다. <표 2-1>과 <표 2-2>에서 중앙부처별 및 지방자치단체별 추출 인원과 함께 이메일 확보 수 및 이의 비율을 제시하였다.

이렇게 확보된 표본수는 중앙부처공무원 1,240명, 지방자치단체 공무원 539명, 연구자 90명으로 총1,869명이었다. 수집된 조사 대상 리스트 1,869명 중 1,125명(60.2%)의 이메일주소를 확보하였다. 이들은 중앙부처에서 735명, 지방자치단체에서 300명, 기타기관에서 90명으로 이루어졌다(<표 2-3>).

<표 2-1> 중앙부처별 추출 인원 및 확보 이메일 수

중앙 부처	추출한 인원 (명)	확보 이메일 수 (%)
감사원	30	30 (100.0)
건설교통부	30	30 (100.0)
경찰청	30	0 (0.0)
공정거래위원회	30	30 (100.0)
국가정보원	40	30 (75.0)
과학기술부	30	30 (100.0)
관세청	40	20 (50.0)
교육인적자원부	30	30 (100.0)
국가보훈처	30	30 (100.0)
국무총리실	30	20 (66.7)
국방부	30	22 (73.3)
국세청	30	0 (0.0)
국정홍보처	30	0 (0.0)
금융감독위원회	30	3 (10.0)
기상청	30	14 (46.7)
기획예산처*	7	0 (0.0)
노동부	40	40 (100.0)
농림부	30	30 (100.0)
농촌진흥청	30	30 (100.0)
문화관광부	30	20 (66.7)
문화재청	30	30 (100.0)
법무부	30	30 (100.0)
법제처	30	0 (0.0)
병무청	40	0 (0.0)
보건복지부	40	10 (25.0)
산림청	30	30 (100.0)
산업자원부	30	16 (53.3)
식품의약품안전청*	12	0 (0.0)
여성가족부	30	30 (100.0)
외교통상부	30	0 (0.0)
재정경제부	40	36 (90.0)
정보통신부	30	15 (50.0)
조달청	30	0 (0.0)
중소기업청	30	0 (0.0)
철도청*	11	4 (36.4)
통일부	30	0 (0.0)
특허청	30	30 (100.0)
해양수산부	30	15 (50.0)
해양경찰청	30	15 (50.0)
행정자치부	30	25 (83.3)
환경부	40	40 (100.0)
합계 (39개 부처)	1,240	735 (59.3)

* 정보공개 30개 미만의 부처

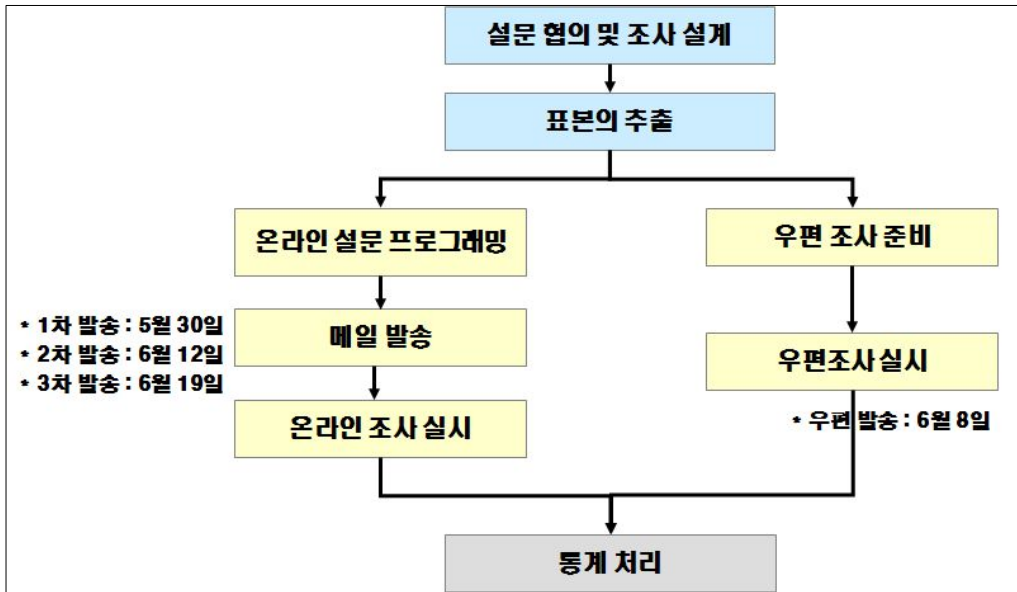
<표 2-2> 지방자치단체별 추출 인원 및 확보 이메일 수

지방자치단체	추출한 인원 (명)	확보 이메일 수 (%)
서울특별시	40	20 (50.0)
부산광역시	30	10 (33.3)
대구광역시	40	37 (92.5)
인천광역시	40	0 (0.0)
광주광역시	40	10 (25.0)
대전광역시	30	20 (66.7)
울산광역시	40	20 (0.0)
경기도	40	40 (100.0)
강원도*	9	7 (77.7)
충청북도	30	15 (50.0)
충청남도	30	20 (66.7)
전라북도	40	40 (100.0)
전라남도	30	0 (0.0)
경상북도	30	15 (50.0)
경상남도	30	18 (60.0)
제주특별자치도	40	28 (70.0)
합계 (16개 시도)	539	300 (55.7)

* 정보공개 30개 미만의 시도

<표 2-3> 조사집단별 이메일 확보비율

구 분	추출한 인원 (명)	확보 이메일 수 (%)
중앙부처	1,240	735 (59.3)
지방자치단체	539	300 (55.7)
대학교수 및 연구기관	90	90 (100.0)
합 계	1,869	1,125 (60.2)



(그림 2-1) 조사과정 요약

가. 온라인조사

온라인조사는 수집된 조사 대상자의 이메일에 조사 안내문((그림 2-2))을 발송하였다. 응답자가 조사 안내문을 통해 조사의 개요와 목적에 대해 이해한 다음, 조사 홈페이지로 이동하여 조사에 참여하는 방식으로 진행되었다. 온라인조사페이지는 설문지의 논리와 설문지의 이동경로에 맞추어 제작하였으며, 응답자의 설문지에 대한 응답내용은 바로 데이터베이스화가 가능하도록 설계되었다((그림 2-3)과 (그림 2-4)).

온라인조사는 2007년 5월 30일부터 6월 22일까지 총 24일 동안 서울시립대 도시사회학과 인터넷조사시스템과 서버를 이용하여 진행되었다. 조사기간 동안 3차례에 걸쳐 이메일을 통해 조사 안내문이 발송하였다. 첫 번째는 5월 30일, 두 번째는 6월 12일, 세 번째는 6월 19일에 발송하였으며, 조사안내 메일의 확인 비율과 조사에 대한 관심도를 높이기 위해 업무에 바쁜 오전시간 대신 오후에 메일을 발송하였다.

안녕하십니까?

조사에 참여해 주셔서 대단히 감사합니다.

통계청은 좀 더 좋은 통계자료를 제공하기 위하여 한국조사연구학회와 공동으로 인구주택총조사 통계자료 이용자 대상의 [인구주택총조사 자료이용실태조사]를 실시하고 있습니다.

선생님의 응답은 인구주택총조사의 효용성을 높이는 방안을 마련하는 데 귀중한 자료가 될 것입니다.

본 조사의 내용은 맞고 틀리는 답이 있는 것이 아닙니다. 선생님께서 평소 생각하신대로 솔직하게 응답해 주시면 됩니다. 응답에 소요되는 시간은 약 10분 정도입니다. 응답해 주신 모든 분들의 개인정보는 통계법 제 13 조에 의해 철저히 보호 받습니다.

응답해 주시는 분들 가운데 20명을 선정하여 5만원 상당의 문화상품권(또는 도서상품권)을 조사 종료 후에 보내드리겠습니다.

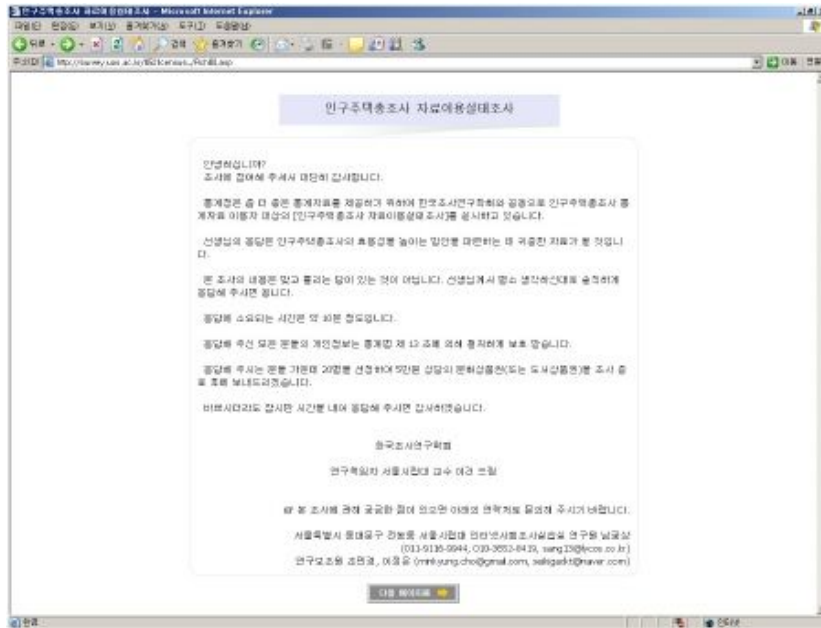
바쁘시더라도 잠시만 시간을 내어 응답해 주시면 감사하겠습니다.

한국조사연구학회
연구책임자 서울시립대 도시사회학과 이건 교수 드림

☞ 본 조사에 관해 궁금한 점이 있으면 아래의 연락처로 문의해 주시기 바랍니다.

서울특별시 동대문구 전농동 서울시립대 인터넷사회조사실습실 연구원 남궁상
(011-9116-9944, 010-3652-8419, sang13@lycos.co.kr)

(그림 2-2) 온라인조사 안내문



(그림 2-3) 웹 서버이 초기화면
(<http://survey.uos.ac.kr/2007survey/index.html>)



(그림 2-4) 웹 서버이 설문화면
(<http://survey.uos.ac.kr/2007survey/index.html>)

응답회수과정을 살펴보면, 1차로 진행된 6월 11일까지 응답한 사람은 93명에 불과하였다. 응답기간이 1.5주 정도 밖에 되지 않았지만, 조사 설계 시에 계획했던 유효표본수 150명을 고려한다면, 응답률은 저조하였다. 이에 아직 설문에 참여하지 않았거나 응답을 끝까지 하지 않고 중간에 설문페이지에서 나간 응답 대상자들에게 조사 참여를 부탁하는 이메일을 발송하였으며, 6월 18일까지의 진행상황을 파악하여 2차 메일 발송 때와 같은 방식으로 6월 19일에 조사 참여를 부탁하는 메일을 다시 한 번 발송하였다.

나. 오프라인조사

오프라인조사는 수집된 1,869명 중 이메일 주소를 확보할 수 없었던 744명에게 준비기간을 거쳐, 6월 8일 우편으로 설문지를 발송하는 것으로 시작하였다. 응답대상자가 배달된 설문에 응답을 한 후, 동봉되어 있는 반송봉투에 넣어 우편으로 회신하는 방법을 사용했으며, 이메일 리스트를 확보할 수 없는 리스트에 대한 온라인조사 대체방법으로 진행되었다. 응답자들에게 설문이 6월 30일까지 도착될 수 있도록 협조를 부탁하였으며, 우편응답은 시일이 걸리기 때문에 7월 5일까지 도착된 설문까지 응답표본수로 간주하였다.

제3절. 응답자의 특성 및 분포

1. 응답률

<표 2-4>는 온라인조사와 오프라인조사(우편조사)의 응답률을 제시한 것이다. 먼저 온라인조사를 살펴보면, 마감일인 6월 22일까지 1,125명의 응답대상자 중 173명이 응답을 하여 15.4%의 응답률을 보였다. 이 중 중앙부처 공무원이 90건(52.0%)이었으며, 지방자치단체 공무원 65건(37.6%), 대학교수 및 기타연구기관이 18건(10.4%)이었다. 각 집단별로 살펴보면, 지방자치단체 공무원들의 응답률은 21.7%로 높았던 반면, 중앙부처 공무원들의 응답률은 12.2%로 낮게 나타났다.

오프라인조사는 7월 5일까지 도착한 설문지 총 301개였으며, 이 중 자료로 사용하기에 부적합한 106개를 제외한 최종 응답수는 195개였다. 이는 오프라인 조사대상 744개 대비 26.2% 응답률이며, 온라인조사의 응답률 15.4% 대비 10.8%p 높은 수치이다. 이 중 중앙부처 공무원이 87건(44.6%)이었으며, 지방자치단체 공무원 108건(55.4%)이었다. 각 집단별 응답률은 중앙부처 공무원들은 17.2%에 그쳤던 반면, 지방자치단체 공무원들이 45.2%로 매우 높게 나타났다.

<표 2-4> 응답대상에 따른 온·오프라인 응답률 분포

	온라인조사			오프라인조사(우편조사)		
	발송 (명)	응답 (명)	응답률 (%)	발송 (명)	응답 (명)	응답률 (%)
전체	1,125	173	15.4	744	195	26.2
중앙부처	735	90	12.2	505	87	17.2
지방자치단체	300	65	21.7	239	108	45.2
기타	90	18	20.0			

흥미로운 사실은 지난 2005년 총조사를 위해 통계청에서는 2003년에 이번 조사와 비슷한 이용자 조사를 했다는 점이다. 그와 비교하기 위해 전체 응답률을 <표 2-5>에 제시하였다. 2003년 조사에서도 온라인과 오프라인 조사를 병행하였다. 다만 오프라인 조사의 경우 이번 조사의 직접 입력 방식과 달리 한글로 작성된 설문지(조사표)를 다운받아 작성하여 이메일로 회신하는 방식을 택했다. 전체적인 응답률은 이번 조사가 19.4%로 2003년 조사의 31.0%에 비해 10% 가량 떨어진다. 그 이전의 2000년 총조사를 위한 이용자 조사에서도 응답률은 31.2%였던 것으로 보아 통계청 조사의 응답률은 매우 안정적이다. 이에 반해 이번 조사는 응답률이 떨어지는 취약점이 있다. 반면 중앙부처 공무원에 대해서는 177명을 조사하여 2003년 조사의 19명보다 훨씬 많은 사람을 조사하였다. 이런 결과는 중앙부처에 대해서는 상대적으로 더 나은 조사결과를 보여주는 강점이 있다.

<표 2-5> 이번 조사와 2003년 조사의 비교

	이번 조사			2003년 조사*		
	발송 (명)	응답 (명)	응답률 (%)	발송 (명)	응답 (명)	응답률 (%)
전체	1,899	368	19.4	728	226	31.0
중앙부처	1,240	177	14.3	72	19	26.4
지방자치단체	539	173	32.1	348	140	40.2
기타	90	18	20.0	308	67	21.8

*통계청 인구조사과(2003)

참고: 2000년 인구주택총조사의 경우 507개 기관 중 158개 응답 (응답률: 31.2%)

2. 응답자의 특성

가. 성별분포

설문 참여자의 성별 분포는 남성이 75.0%로 매우 높은 비율을 차지했다. 중앙부처 소속의 남성비율이 78.5%로 지방자치단체 소속 남성비율 72.3%에 비해 높게 나타났으며, 연구소나 대학교수 등이 소속된 기타기관의 응답자 분포에서는 남성의 비율이 66.7%로 다소 낮은 편이었다. 하지만, 기타기관의 경우 총 응답자수가 18명밖에 되지 않는 관계로 다른 집단과의 차이를 말할 수는 없다. <표 2-6>은 각 소속기관별로 성별 분포를 제시한 것이다.

<표 2-6> 응답자의 성별 분포

소속기관 \ 성별	전체		중앙부처		지방자치단체		기타기관	
	명	(%)	명	(%)	명	(%)	명	(%)
남 자	276	(75.0)	139	(78.5)	125	(72.3)	12	(66.7)
여 자	92	(25.0)	38	(21.5)	48	(27.7)	6	(33.3)
계	368	(100.0)	177	(100.0)	173	(100.0)	18	(100.0)

나. 연령별 분포

연령대 분포를 살펴보면, <표 2-7>과 같다. 20대의 비율은 4.9%로 매우 낮은 비율인 반면, 40대는 42.4%, 50대 이상 응답자도 24.2%로 나타나 전반적인 연령대는 높은 편이었다. 특히, 지방자치단체 소속응답자들 중 50대 이상이 29.5%로 중앙부처나 기타기관에 비해 더 높은 연령층을 형성하고 있었다.

<표 2-7> 응답자의 연령 분포

연령 \ 소속기관	전체		중앙부처		지방자치단체		기타기관	
	명	(%)	명	(%)	명	(%)	명	(%)
20대	18	(4.9)	9	(5.1)	8	(4.6)	1	(5.6)
30대	105	(28.5)	56	(31.6)	41	(23.7)	8	(44.4)
40대	156	(42.4)	77	(43.5)	73	(42.2)	6	(33.3)
50대 이상	89	(24.2)	35	(19.8)	51	(29.5)	3	(16.7)
계	368	(100.0)	177	(100.0)	173	(100.0)	18	(100.0)

조사방법에 따른 연령대 분포 비교는 <표 2-8>과 같다. 20대에서 온라인조사 응답자의 비율이 약간 높을 뿐, 별 차이를 보이지 않았다.

<표 2-8> 조사방법별 응답자의 연령 분포

연령 \ 조사방법	전체		온라인		오프라인	
	명	(%)	명	(%)	명	(%)
20대	18	(4.9)	13	(7.5)	5	(2.6)
30대	105	(28.5)	47	(27.2)	58	(29.7)
40대	156	(42.4)	69	(39.9)	87	(44.6)
50대 이상	89	(24.2)	44	(25.4)	45	(23.1)
계	368	(100.0)	173	(100.0)	195	(100.0)

했던 병무청의 참여빈도가 28명으로 가장 높았으며, 다음으로 관세청이 13명의 높은 참여빈도를 보였다.

<표 2-9> 중앙부처별 발송건수 및 응답률

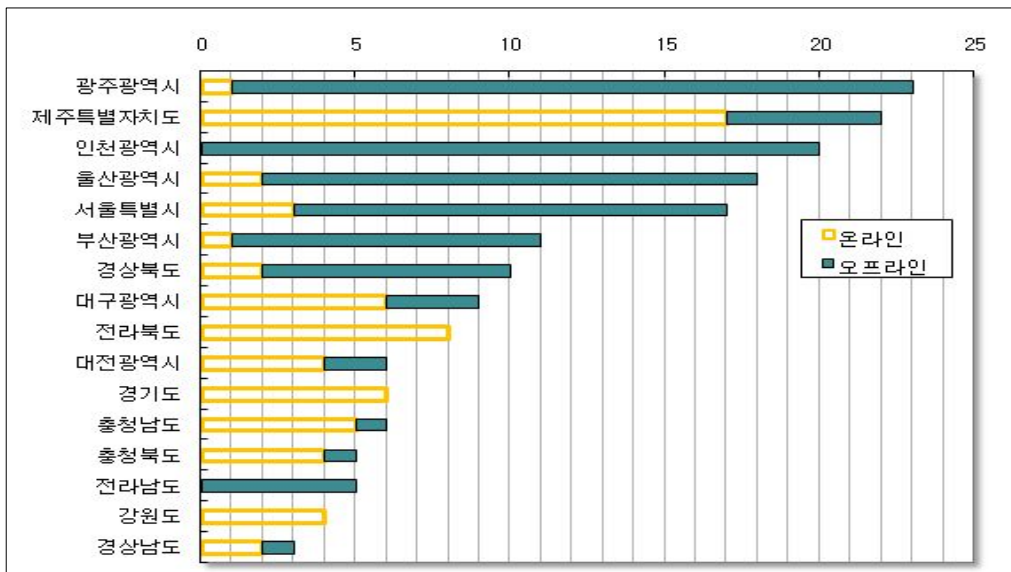
구 분	발송건수(명)	응답자수(명)	응답률(%)
건설교통부	30	9	30.0
국방부	30	3	10.0
과학기술부	30	2	6.7
노동부	40	3	7.5
농림부	30	2	6.7
문화관광부	30	5	16.7
보건복지부	40	11	27.5
산업자원부	30	4	13.3
정보통신부	30	6	20.0
재정경제부	40	1	2.5
통일부	30	3	10.0
해양수산부	30	1	3.3
여성가족부	30	2	6.7
국세청	30	2	6.7
국정홍보처	30	3	10.0
관세청	40	14	35.0
기상청	30	6	20.0
농촌진흥청	30	3	10.0
법제처	30	5	16.7
병무청	40	28	70.0
산림청	30	1	3.3
식품의약품안전청	12	8	66.7
조달청	30	7	23.3
중소기업청	30	7	23.3
특허청	30	4	13.3
해양경찰청	30	2	6.7
감사원	30	16	53.3
공정거래위원회	30	2	6.7
국가정보원	40	10	25.0
국무총리실	30	5	16.7
금융감독위원회	30	2	6.7
합 계	1,240	177	14.3

발송건수 대비 응답률에서는 오프라인 조사 참여가 가장 높았던 병무청이 60.7%로 가장 높은 수치를 나타냈다(<표 2-9>). 병무청 다음으로는 발송건

수 12명 가운데 8명이 응답한 식품의약품안전청(66.7%)이었으며, 감사원 역시 53.3%의 높은 참여율을 보였다.

(2) 지방자치단체

지방자치단체 중 조사에 가장 적극적인 태도를 보인 단체는 광주광역시와 제주특별자치도였다((그림 2-6)). 광주광역시는 총 23명이, 제주특별자치도는 22명이 조사에 참여했으며, 다음으로 인천광역시가 20명, 울산광역시가 18명, 서울특별시가 17명으로 높은 참여율을 보였다. 반면, 경상남도는 3명, 강원도는 4명이 참여하는 데 그쳐 가장 낮은 참여율을 나타냈다. 이밖에 충청북도, 전라남도, 대전광역시 등도 그리 높지 않은 조사 참여율을 보였다.



(그림 2-6) 지방자치단체별 응답자수

조사방법별로 살펴보면, 온라인조사에서는 제주특별자치도가 17명으로 가장 높은 참여빈도를 보였으며, 다음으로는 전라북도가 8명의 참여빈도를 보였다. 반면, 오프라인조사의 참여빈도를 살펴보면, 온라인조사에서 낮았던 광주광역시의 참여빈도가 22명으로 가장 높았으며, 다음으로 온라인조사를 실시하지 않았던 인천광역시가 20명의 높은 참여빈도를 보였다.

발송건수 대비 응답률에서도 광주광역시와 제주특별자치도가 각각 57.5%와 55.0%로 가장 높은 수치를 나타냈다(<표 2-10>). 다음으로는 인천광역시(50.0%)였으며, 울산광역시가 45.0%, 강원도가 44.4%의 응답률을 보였다. 이에 비해 경기도, 경상남도, 충청북도, 전라남도 등은 10%의 가장 저조한 응답률을 나타냈다.

<표 2-10> 지방자치단체별 발송건수 및 응답률

구 분	발송건수(명)	응답자수(명)	응답률(%)
서울특별시	40	17	42.5
부산광역시	30	11	36.7
대구광역시	40	9	22.5
인천광역시	40	20	50.0
광주광역시	40	23	57.5
대전광역시	30	6	20.0
울산광역시	40	18	45.0
경기도	40	6	15.0
강원도	9	4	44.4
충청북도	30	5	16.7
충청남도	30	6	20.0
전라북도	40	8	20.0
전라남도	30	5	16.7
경상북도	30	10	33.3
경상남도	30	3	10.0
제주특별자치도	40	22	55.0
합 계	539	173	32.1

라. 직급 분포

직급별 분포를 살펴보면, <표 2-11>과 같다. 연구기관, 대학교수 등이 속해 있는 기타기관 응답자들은 공무원 직급체계에 의해 분류하기에는 한계가 있어 직급 분석에서는 제외하였다. 응답자의 직급은 3급에서 9급까지와 계약직 등의 분포를 이루고 있었다. 5급과 6급이 각각 95명(27.1%), 102명(29.1%)으로 가장 높은 비중을 차지하였으며, 다음으로 7급이 66명(18.9%)이었다. 중앙부처의 경우, 5급이 57명(32.2%)으로 가장 많았으며, 6급이 42명(23.7%), 4

급 34명(19.2%)으로 그 뒤를 이었다. 반면, 지방자치단체는 6급이 60명(34.7%)으로 가장 많았고, 다음으로 7급이 40명(23.1%), 5급이 38명(22.0%)이었다.

<표 2-11> 응답자의 소속기관별 직급 분포

소속기관 직급	전체		중앙부처		지방자치단체	
	명	(%)	명	(%)	명	(%)
3급	4	(1.1)	4	(2.3)	0	(0.0)
4급	38	(10.9)	34	(19.2)	4	(2.3)
5급	95	(27.1)	57	(32.2)	38	(22.0)
6급	102	(29.1)	42	(23.7)	60	(34.7)
7급	66	(18.9)	26	(14.7)	40	(23.1)
8급	18	(5.1)	9	(5.1)	9	(5.2)
9급	12	(3.4)	4	(2.3)	8	(4.6)
10급	7	(2.0)	0	(0.0)	7	(4.0)
계약직	8	(2.3)	1	(0.6)	7	(4.0)
계	350	(100.0)	177	(100.0)	173	(100.0)

제4절. 인구주택총조사 자료 이용 현황

1. 일반적인 통계 자료 이용

인구주택총조사 자료 이용에 앞서 일반적으로 얼마나 통계를 자주 이용하는지 알아보았다. <표 2-12>는 그에 대한 분포이다. 응답자의 18.5%가 “매우 자주 이용한다”라고 응답했고, 45.7%는 “자주 이용하는 편”이라고 응답해 전체 응답자의 64.2%가 통계자료를 업무에 자주 이용하고 있는 것으로 나타났다. 이는 공무원들이나 연구원들이 일반 통계자료를 업무에 자주 이용하는 편이라고 할 수 있다. 기관별로 자주 이용하는 비율을 살펴보면, 중앙부처 공무원은 58.8%(매우 자주 이용한다 16.4% + 자주 이용하는 편 42.4%), 지방자치단체공무원 67.1%, 기타기관 소속 응답자 88.9%로 나타나, 중앙부처와

지방자치단체 공무원보다는 기타기관 소속 응답자가 통계자료 이용률에서 앞서는 것으로 나타났다. 반면, 전체 응답자의 4.9%(중앙부처 6.8%, 지자체 3.5%)는 통계자료를 전혀 업무에 이용하지 않고 있었다.¹⁾

<표 2-12> 소속기관별 통계 자료 이용 빈도

이용빈도	소속기관		전체		중앙부처		지방자치단체		기타기관	
	명	(%)	명	(%)	명	(%)	명	(%)	명	(%)
매우 자주 이용한다	68	(18.5)	29	(16.4)	28	(16.2)	11	(61.1)		
자주 이용하는 편	168	(45.7)	75	(42.4)	88	(50.9)	5	(27.8)		
별로 이용하지 않는 편	114	(31.0)	61	(34.5)	51	(29.5)	2	(11.1)		
거의 이용하지 않는다	18	(4.9)	12	(6.8)	6	(3.5)	0	(0.0)		
계	368	(100.0)	177	(100.0)	173	(100.0)	18	(100.0)		
$\chi^2(6)=27.3$ (p-값=0.000)										

<표 2-13>는 통계자료 이용 빈도에 대한 조사방법별 결과를 비교한 표이다. 온라인 응답자(74.6%)가 오프라인 응답자(54.8%)에 비해 통계자료를 자주 이용하는 것으로 나타났다. 이용강도 비교에서도 온라인 응답자들의 30.1%가 통계자료를 “매우 자주 이용 한다”고 응답해, 오프라인 응답자(8.1%)에 비해 훨씬 더 높은 이용강도를 나타냈다.

<표 2-13> 조사방법별 통계 자료 이용 빈도

이용빈도	소속기관		전체		온라인		오프라인	
	명	(%)	명	(%)	명	(%)	명	(%)
매우 자주 이용한다	68	(18.5)	52	(30.1)	16	(8.1)		
자주 이용하는 편	168	(45.7)	77	(44.5)	91	(46.7)		
별로 이용하지 않는 편	114	(31.0)	37	(21.4)	77	(39.5)		
거의 이용하지 않는다	18	(4.9)	7	(4.0)	11	(5.6)		
계	368	(100.0)	173	(100.0)	195	(100.0)		
$\chi^2(3)=34.0$ (p-값=0.000)								

1) 한가지 유의할 점은 상대적으로 많은 사람들이 통계를 이용하는 것으로 응답하였는데, 이것은 어쩌면 통계를 다루는 사람들이 주로 응답을 한 때문에 나타난 것일 수도 있다.

2. 인구주택총조사 자료 이용

가. 인구주택총조사 자료 이용경험

인구주택총조사 자료 이용경험률은 75.8%로 나타났다(<표 2-14>). 지방자치단체와 기타기관의 이용경험이 각각 84.4%, 94.4%으로 높았던 반면, 중앙부처는 65.6%로 상대적으로 낮은 이용경험률을 보였다. 특히, 주된 관심인 중앙부처와 지방자치단체의 차이는 20%p 가량으로 많은 차이를 보였으며, 그러한 차이는 1%의 유의수준에서 통계적으로 의미가 있었다($\chi^2(2)=16.6$, P-값=0.000). 이것은 지방자치단체가 중앙부처보다 인구주택총조사를 더 많이 이용하는 것을 보여준다. 이러한 이유는 인구주택총조사 자료가 지방자치단체 업무에 더 유용하기 때문이며, 더욱이 2005년 조사에서 시도수준의 16개 지방자치단체별로 3개의 항목들을 추가하여 조사하여 이런 경향이 강화됐기 때문이기도 하다.

<표 2-14> 소속기관별 인구주택총조사 자료 이용경험

이용여부 \ 소속기관	전체		중앙부처		지방자치단체		기타기관	
	명	(%)	명	(%)	명	(%)	명	(%)
이용 경험 있음	279	(75.8)	116	(65.6)	146	(84.4)	17	(94.4)
이용 경험 없음	89	(24.2)	61	(34.4)	27	(15.6)	1	(5.6)
계	368	(100.0)	177	(100.0)	173	(100.0)	18	(100.0)
		$\chi^2(2)=20.5$ (p-값=0.000)						

조사방법별로 살펴보면, 온라인조사 응답자들의 이용경험률은 87.3%로 오프라인조사(65.6%)에 비해 높게 나타났다 (<표 2-15>). 이는 앞서 제시한

통계자료 이용 빈도 결과(<표 2-13>)에서 온라인조사 응답자들이 오프라인 조사 응답자들에 비해 더 높은 이용 빈도를 나타낸 것과 관계가 있다.

<표 2-15> 조사방법별 인구주택총조사 자료 이용경험

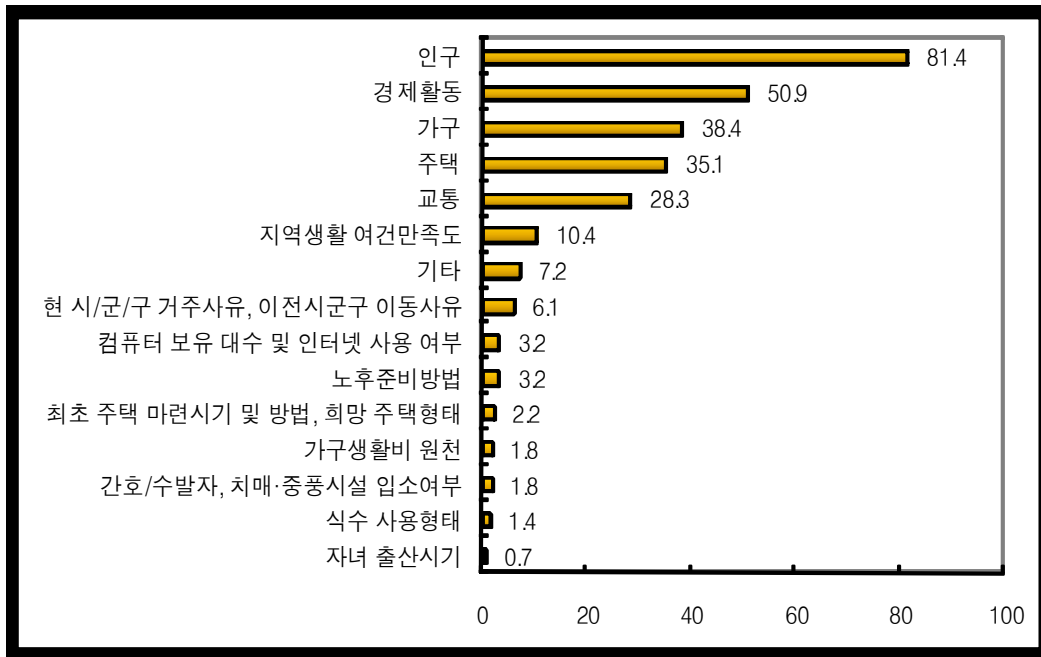
소속기관 \ 이용빈도	전체		온라인		오프라인	
	명	(%)	명	(%)	명	(%)
이용 경험 있음	279	(75.8)	151	(87.3)	128	(65.6)
이용 경험 없음	89	(24.2)	22	(12.7)	67	(34.4)
계	368	(100.0)	173	(100.0)	195	(100.0)
			$\chi^2(1)=23.4$ (p-값=0.000)			

나. 인구주택총조사 자료별 이용경험률

인구주택총조사 자료별 이용경험률은 (그림 2-7)과 같다. 이용률이 가장 높은 자료는 인구부문으로 인구주택총조사 자료이용경험자의 81.4%로 나타났다. 다음으로는 경제활동 부문이 50.9%의 이용경험률을 보였으며, 가구부문 38.4%, 주택부문 35.1%, 교통부문 28.3%의 이용경험률을 나타냈다.

중앙부처 공무원은 인구주택총조사 자료 이용경험과 마찬가지로 세부자료별 이용경험률이 다른 기관에 비해 낮은 편이었다(<표 2-16>). 전반적으로 가장 높은 이용률을 보였던 인구부문에서 지방자치단체나 기타기관은 90%에 가까운 이용경험률을 보인 반면, 중앙부처 소속 공무원들은 69.8%에 그쳤다. 중앙부처 공무원은 경제활동 부문만 49.1%로 지방자치단체(48.6%)보다 높고, 다른 항목 이용경험률은 지방자치단체에 비해 다소 낮은 편이었다.

지방자치단체는 경제활동, 가구, 주택 관련 자료 등에서 고른 이용경험률(48.6%)을 보였다. 또한, 2005년도 총조사에서 지방자치단체별 추가조사항목이었던 지역생활 여건만족도(11.9%), 현 거주지 거주사유 및 이전사유(11.6%)도 높은 이용경험률을 보였다.



(그림 2-7) 인구주택총조사 자료별 이용경험률 (복수응답)

<표 2-16> 인구주택총조사 자료별 이용경험률 (복수응답)

이용자료	소속기관		중양부처		지자체		기타기관	
	명	(%)	명	(%)	명	(%)	명	(%)
인구	227	(81.4)	81	(69.8)	131	(89.7)	15	(88.2)
경제활동	142	(50.9)	57	(49.1)	71	(48.6)	14	(82.4)
가구	107	(38.4)	26	(22.4)	71	(48.6)	10	(58.8)
주택	98	(35.1)	21	(18.1)	68	(46.6)	9	(52.9)
교통	79	(28.3)	21	(18.1)	51	(34.9)	7	(41.2)
기타	20	(7.2)	6	(5.2)	13	(8.9)	1	(5.9)
자녀 출산시기	2	(0.7)			2	(1.4)		
현 거주지거주사유/이전사유	17	(6.1)			17	(11.6)		
노후준비방법	9	(3.2)			9	(6.2)		
간호수발자/치매중풍시설입소	5	(1.8)			5	(3.4)		
최초주택마련시기/방법	6	(2.2)			6	(4.1)		
식수 사용형태	4	(1.4)			4	(2.7)		
지역생활 여건만족도	29	(10.4)			29	(19.9)		
컴퓨터보유/인터넷사용	9	(3.2)			9	(6.2)		
가구생활비 원천	5	(1.8)			5	(3.4)		

<표 2-16>의 항목 중 인구/교통/경제활동/가구/주택/기타 부문을 제외한 자녀출산시기, 현 시군구 거주사유 등의 항목들은 시도별 표본조사 항목으로 구성된 자료들로서, 지방자치단체들마다 다른 조사항목을 조사되었다. 참고로 시도별 표본조사 항목은 <표 2-17>에 제시하였다. 이들은 시도별로 다르게 조사된 추가구성 항목으로 공통항목에 비해 전체적인 이용률은 낮다. 이 가운데 상대적으로 높은 이용률(19.9%)을 보인 지역생활 여건 만족도는 8개 시도에서, 그 다음으로 높은 이용률(11.6%)을 보인 현 거지지 거주사유 및 이전사유는 7개 시도에서 제안한 항목으로 여러 지방자치단체에서 관심을 갖는 것이다. 이에 반해 노후준비방법이나 최초주택마련시기 및 방법은 각각 7개와 6개 시도에서 제안했음에도 이용률은 각각 6.2%와 4.1%로 그다지 높지 않은 편이다. 그렇지만 전체적으로 표본의 수가 그다지 많지 않아 이의 통계적 신뢰성은 이 조사로만은 확인하기 어렵다. 앞으로 조사에서는 좀 더 많은 표본수를 확보해야 이러한 항목들의 이용 현황을 좀 더 정확히 파악할 수 있을 것이다.

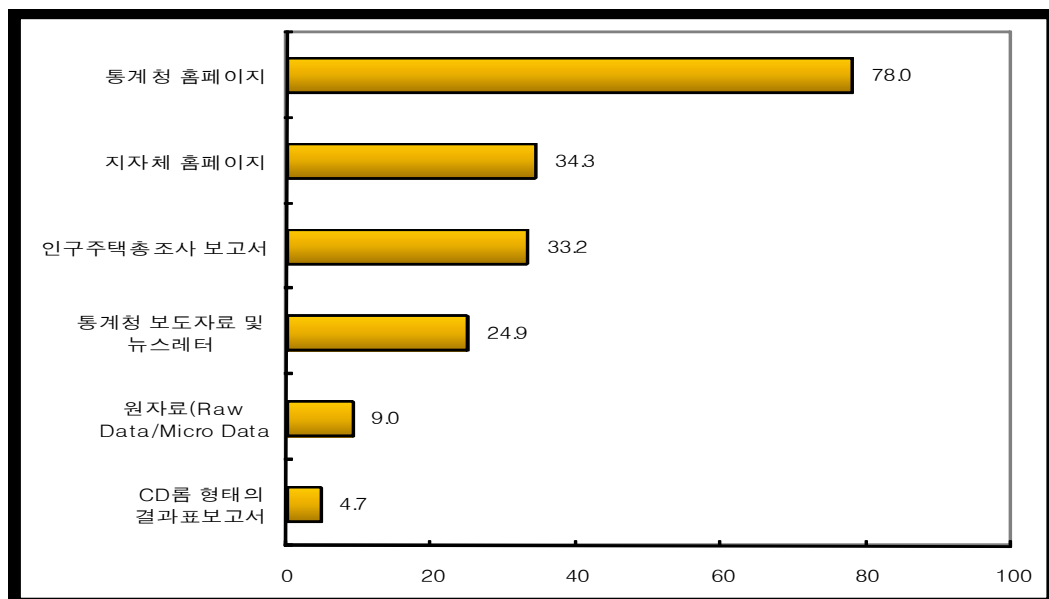
<표 2-17> 시도별 추가표본조사항목 및 표본수

시도별 추가조사항목	표본수	조사 지역
지역생활여건만족도	103	인천, 광주, 대전, 울산, 충남, 전남, 경남, 제주
현거주지거주사유/이전사유	73	서울, 광주, 경기도, 강원도, 전남, 전북, 경북
최초주택마련시기/방법	57	부산, 대구, 인천, 대전, 충북, 충남
노후준비방법	49	부산, 경기도, 강원도, 충북, 전북, 전남, 경북
컴퓨터보유/인터넷사용	46	인천, 강원도, 제주도
간호수발자/치매중풍시설입소	44	부산, 대전, 울산, 충남, 경남
가구생활비원천	32	경북, 제주도
자녀출산시기	26	서울, 경기도, 경남
식수사용형태	14	대구, 충북

마지막으로, 기타기관은 전반적으로 인구주택총조사의 자료 이용률이 높게 나타났다. 그 중 인구부문 자료 이용률이 88.2%로 가장 높았으며, 다음으로 경제활동(82.4%), 가구(58.8%), 주택(52.9%)의 순으로 자료 이용률이 높았다. 이러한 모습은 인구, 경제활동, 가구, 주택에 대한 연구자들의 관심이 높은 것을 나타내며, 더 나아가 우리 사회 전반에서 중요하게 생각하는 사안들을 표시하기도 한다.

다. 인구주택총조사 자료이용 형태

통계청 홈페이지를 통해 이용했다는 응답이 78.0%로 가장 높았다((그림 2-8)). 다음으로 지자체 홈페이지(34.3%), 인구주택총조사 보고서(33.2%), 통계청 보도자료 및 뉴스레터(24.9%), 원자료 (microdata)(9.0%)의 순서로 높게 나타났다.



(그림 2-8) 인구주택총조사 자료이용 형태(복수응답)

이와 같은 현상은 앞으로 통계청이 전통적 매체보다 인터넷을 통한 자료 제공에 더욱 신경을 쓸 필요가 있음을 시사한다. 이때 지방자치단체의 인터

넷을 통한 통계 서비스에도 통계청과 같은 수준의 품질과 같은 자료를 일관되게 제공하도록 지방자치단체에 대한 통계제공지침이나 교육을 강화하고, 그들과 협력관계를 잘 유지할 필요도 함께 시사하고 있다.

(1) 소속기관별 자료이용 형태

자료이용 형태를 소속기관별로 살펴보면 <표 2-18>과 같다. 중앙부처 공무원은 통계청 홈페이지 이용이 79.3%로 가장 높고, 다음으로 통계청 보도자료 및 뉴스레터가 25.0%, 인구주택총조사 보고서가 24.1%의 순이었다. 지자체 공무원은 통계청 홈페이지 이용이 75.0%, 지자체 홈페이지가 52.8%, 인구주택총조사 보고서가 39.6%의 순이었다. 기타기관 응답자는 통계청 홈페이지가 94.1%, 원자료가 64.7%, CD롬 형태 결과표보고서가 47.1%, 인구주택총조사 보고서가 41.2%의 순으로 이용한 것으로 나타났다. 한가지 부연할 것은 기관들의 원자료 이용과 CD이용이 중앙부처나 지방자치단체에 비해 월등히 높다는 점이다. 이것은 이들이 자료를 심층적으로 다루고, 자신들에게 필요한 통계를 가공하기 때문이다. 이러한 추세에 부응하여 통계청은 총조사 자료의 활용성을 높이기 위해 원자료 서비스와 함께 전자적으로 활용할 수 있는 자료제공방법을 지속적으로 확대할 필요가 있다.

<표 2-18> 인구주택총조사 이용 자료형태 (복수응답)

관	소속기	전체	중앙부처	지방자치단체	기타기관
		(n=279)	(n=116)	(n=146)	(n=17)
이용 자료형태		명 (%)	명 (%)	명 (%)	명 (%)
통계청 홈페이지		216 (78.0)	92 (79.3)	108 (75.0)	16 (94.1)
지자체 홈페이지		95 (34.3)	16 (13.8)	76 (52.8)	7 (17.6)
인구주택총조사 보고서		92 (33.2)	28 (24.1)	57 (39.6)	3 (41.2)
통계정보도자료/뉴스레터		69 (24.9)	29 (25.0)	35 (24.3)	5 (29.4)
원자료		25 (9.0)	5 (4.3)	9 (6.3)	11 (64.7)
CD롬 형태 결과표보고서		13 (4.1)	3 (2.6)	2 (1.4)	8 (47.1)

(2) 조사방법별 자료이용 형태

온라인조사와 오프라인조사 응답자들 모두 통계청 홈페이지에서 제공하는 자료를 주로 이용하고 있는 것으로 나타났다<표 2-19>. 그 다음으로 오프라인조사에서는 지자체 홈페이지, 인구주택총조사 보고서의 순서로 이용하는 비율이 높게 나타난 반면, 온라인조사에서는 인구주택총조사 보고서, 지방자치단체 홈페이지의 순서로 이용하고 있었다. 이것은 지방자치단체 공무원들이 지자체 홈페이지를 많이 이용하는 데 이들이 또한 오프라인조사에 많이 응답한 때문이기도 하다(<표 2-4> 참조). 또한 온라인조사 응답자들은 원자료의 이용비율도 15.2%로 나타나, 오프라인조사 응답자들(1.6%)에 비해 높은 이용률을 보였다.

<표 2-19> 조사방법별 인구주택총조사 이용 자료형태 (복수응답)

소속기관 이용빈도	전체 (n=279)		온라인 (n=151)		오프라인 (n=128)	
	명	(%)	명	(%)	명	(%)
통계청 홈페이지	216	(78.0)	118	(78.1)	98	(77.8)
지자체 홈페이지	95	(34.3)	41	(27.2)	54	(42.9)
인구주택총조사 보고서	92	(33.2)	51	(33.8)	41	(32.5)
통계정보도자료/뉴스레터	69	(24.9)	31	(20.5)	38	(30.2)
원자료	25	(9.0)	23	(15.2)	2	(1.6)
CD롬 형태 결과표보고서	13	(4.1)	11	(7.3)	2	(1.6)

라. 인구주택총조사 자료이용 용도

인구주택총조사 자료의 주된 이용용도에서는 정책수립 및 평가를 목적으로 이용하는 비중이 62.9%로 가장 높았으며, 연구·학술·학습활동을 위한 이

용은 21.9%였다(<표 2-20>). 정책수립 및 평가와 연구·학술·학습활동을 합한 비중이 84.8%로 나타나, 자료 이용의 대부분이 이들 두 부문으로 한정되었다.

집단별로 용도의 차이가 나는데, 중앙부처와 지방자치단체들은 정책수립 및 평가를 위한 이용비중이 높았던 반면, 기타기관에서는 연구·학술·학습활동을 위해 자료를 사용하는 경우가 많았다. 다시 말해 공무원은 정책, 연구소와 학계는 연구와 학술활동에서 인구주택총조사를 이용하여 기관의 업무 성격을 잘 반영하였다.

<표 2-20> 인구주택총조사 자료의 주된 이용용도 (복수응답)

주이용용도 \ 소속기관	전체 (n=279)		중앙부처 (n=116)		지방자치단체 (n=146)		기타기관 (n=17)	
	명	(%)	명	(%)	명	(%)	명	(%)
정책 수립 및 평가	175	(62.9)	78	(67.2)	97	(66.9)	0	(0.0)
연구·학술·학습활동	61	(21.9)	24	(20.7)	24	(16.6)	13	(76.5)
사업 및 경영계획 수립	41	(14.7)	7	(6.0)	34	(23.4)	0	(0.0)
교육자료 인용	37	(13.3)	7	(6.0)	26	(17.9)	4	(23.5)
표본조사의 모집단	12	(4.3)	6	(5.2)	6	(4.1)	0	(0.0)
기타	16	(5.8)	5	(4.3)	11	(7.6)	0	(0.0)

마. 인구주택총조사 자료이용 빈도

응답자의 11.2%가 “매우 자주 이용한다”라고 응답했고, 44.2%는 “자주 이용하는 편”이라고 응답해 전체 응답자의 55.4%가 인구주택총조사 자료를 업무에 자주 이용하고 있는 것으로 나타났다(<표 2-21>). 자주 이용하는 비율은 소속기관별로 차이가 있었다(표의 χ^2 참고). 이를 좀 더 상세히 보기 위해 ‘매우 자주 이용’과 ‘자주 이용하는 편’을 합쳐 Heavy User로, ‘별로

이용하지 않는 편'과 '거의 이용하지 않는다'를 합쳐 Light User로 나누어 보았다. 기관별로 Heavy User의 비율이 중앙부처 공무원은 47%, 지방자치단체 공무원은 58.9%, 기타기관 응답자는 82.4%이어서 기타기관의 이용자가 월등히 많은 것으로 나타났다. 주요 관심사인 중앙부처와 지방자치단체만을 보면 그 차이가 12% 정도이며, 1%의 유의수준에서 통계적으로 의미가 있었다($\chi^2(2)=9.0$, p-값=0.001). 이런 결과는 중앙부처보다 지방자치단체에서 인구주택총조사를 더 많이 이용하는 것을 보여준다.

<표 2-21> 인구주택총조사 자료이용 빈도

이용빈도		소속기관		전체		중앙부처		지방자치단체		기타기관	
		명	(%)	명	(%)	명	(%)	명	(%)	명	(%)
Heavy User	매우 자주 이용한다	31	(11.2)	11	(9.6)	11	(7.5)	9	(53.0)		
	자주 이용하는 편	123	(44.2)	43	(37.4)	75	(51.4)	5	(29.4)		
Light User	별로 이용하지 않는 편	113	(40.6)	51	(44.3)	59	(40.4)	3	(17.6)		
	거의 이용하지 않는다	11	(4.0)	10	(8.7)	1	(0.7)	0	(0.0)		
계		278	(100.0)	115	(100.0)	146	(100.0)	17	(100.0)		
				$\chi^2(6)=46.1$ (p-값=0.000)							

<표 2-22>는 인구주택총조사 자료 이용 빈도에 대한 조사방법별 결과를 비교한 표이다. 온라인 응답자(64.3%)가 오프라인 응답자(44.9%)에 비해 통계자료를 자주 이용하는 것으로 나타났다. 이용강도 비교에서도 온라인 응답자들의 16.6%가 통계자료를 “매우 자주 이용 한다”고 응답해, 오프라인 응답자(4.7%)에 비해 더 강한 이용강도를 보이고 있었다.

<표 2-22> 조사방법별 인구주택총조사 자료 이용 빈도

이용빈도 \ 소속기관	전체		온라인		오프라인	
	명	(%)	명	(%)	명	(%)
매우 자주 이용한다	31	(11.2)	25	(16.6)	6	(4.7)
자주 이용하는 편	123	(44.2)	72	(47.7)	51	(40.2)
별로 이용하지 않는 편	113	(40.6)	47	(31.1)	66	(52.0)
거의 이용하지 않는다	11	(4.0)	7	(4.6)	4	(3.1)
계	278	(100.0)	151	(100.0)	127	(100.0)
			$\chi^2(3)=17.3$ (p-값=0.001)			

바. 인구주택총조사 자료 활용의 업무 영향력 평가

(1) 소속기관별 평가

<표 2-23>은 인구주택총조사 자료 활용의 업무 영향력을 평가한 결과이다. 응답자의 21.7%가 매우 중요하다고 응답했고, 44.8%는 중요한 편이라고 응답했다. 전체 응답자의 65% 이상이 인구주택총조사 자료 활용이 업무에 중요한 영향을 주는 것으로 인식하고 있었으며, 소속기관별로 이러한 인식에는 차이가 거의 없었다.

(2) 자료이용강도별 평가

Heavy User와 Light User 두 집단이 인구주택총조사를 어떻게 다르게 평가하는지 살펴보았다(<표 2-24>). 이용강도가 강한 Heavy User는 인구주택총조사 자료가 본인의 업무에 중요하게 활용된다고 생각하는 비율이 82.5%(매우 중요하다 29.9%, 중요한 편이다 52.6%)로 높게 나타난 반면, Light User는 중요하다는 인식이 45.9%에 그쳐 그 차이가 크게 나타났다. 이러한 평가의 차이는 역시 인구주택총조사가 실제 업무에서 더욱 많이 이용될수록 그에 대한 중요성의 인식도 높아진다는 것을 의미하며, 더 나

아가 인구주택총조사의 활용성을 높이는 것이 총조사의 중요성을 인식하게 하는 하나의 방안임을 보여주고 있다.

<표 2-23> 인구주택총조사 자료 활용의 업무 영향력 평가

이용빈도	소속기관		중양부처		지방자치단체		기타기관	
	명	(%)	명	(%)	명	(%)	명	(%)
매우 중요하다	60	(21.7)	23	(19.8)	25	(17.4)	12	(70.6)
중요한 편이다	124	(44.8)	50	(43.1)	72	(50.0)	2	(11.8)
보통이다	75	(27.1)	34	(29.3)	39	(27.1)	2	(11.8)
중요하지 않은 편이다	18	(6.5)	9	(7.8)	8	(5.6)	1	(5.9)
전혀 중요하지 않다	0	(0.0)	0	(0.0)	0	(0.0)	0	(0.0)
계	277	(100.0)	116	(100.0)	144	(100.0)	17	(100.0)
$\chi^2(6)=27.5$ (p-값=0.000)								

<표 2-24> 자료이용강도별 평가

영향력	이용강도		Heavy User		Light User	
	명	(%)	명	(%)	명	(%)
매우 중요하다	60	(21.7)	46	(29.9)	14	(11.5)
중요한 편이다	123	(44.6)	81	(52.6)	42	(34.4)
보통이다	75	(27.2)	27	(17.5)	48	(39.3)
중요하지 않은 편이다	18	(6.5)	0	(0.0)	18	(14.8)
전혀 중요하지 않다	0	(0.0)	0	(0.0)	0	(0.0)
계	276	(100.0)	154	(100.0)	122	(100.0)
$\chi^2(3)=50.3$ (p-값=0.000)						

제5절. 인구주택총조사 자료 이용 만족도

1. 전반적 만족도

인구주택총조사 자료 이용경험을 바탕으로 만족도를 측정한 결과는 <표 2-25>와 같다. 먼저 전체적으로 보면, 인구주택총조사 자료에 대해 매우 만족하는 비율은 3.2%였으며, 약간 만족하는 비율은 55.8%였다. 매우 만족과 약간 만족을 합한 만족비율은 59.0%였다. 소속기관별로 매우 만족과 약간 만족을 합하여 만족비율을 살펴보면, 기타기관이 70.6%로 가장 높았다. 중앙부처와 지방자치단체는 대체로 비슷한 분포를 보였는데, 그 가운데도 지방자치단체가 조금 더 만족한 모습을 나타냈다. 중앙부처의 만족비율은 53.9%, 지방자치단체는 61.6%로 중앙부처가 가장 낮은 만족비율을 나타냈으나 5%의 유의수준에서 통계적으로 의미 있는 것은 아니었다($\chi^2(3)=1.64$, p-값=0.651). 5점 척도 평균에서도 중앙부처(3.54점) 보다는 지방자치단체(3.62점)가 높은 점수를 기록했으나 이 역시 통계적으로 의미 있는 결과는 아니다($t=1.104$, p-값=0.271).

한 가지 재미있는 논의는 마케팅연구에서 찾아 볼 수 있다. Zeithaml and Bitner(1996)는 만족도 조사에서 60-70% 정도의 만족수준이 나타나야 이용자가 큰 불만 없이 이용하는 수준이라고 보았다. 이들의 논의를 따르면, 중앙부처는 인구주택총조사에 만족하지 못하는 편이며, 지방자치단체는 어느 정도 만족하는 편이라고 할 수 있다. 이렇듯 조금 어긋나는 결과는 이에 관한 좀 더 심도 있는 연구의 필요성을 제시한다.

<표 2-25> 인구주택총조사 자료에 대한 전반적 만족도

전반적 만족도	소속기관		중양부처		지방자치단체		기타기관	
	명	(%)	명	(%)	명	(%)	명	(%)
매우 만족	9	(3.2)	4	(3.5)	5	(3.4)	0	(0.0)
약간 만족	155	(55.8)	58	(50.4)	85	(58.2)	12	(70.6)
보통	104	(37.4)	49	(42.6)	52	(35.6)	3	(17.6)
약간 불만족	10	(3.6)	4	(3.5)	4	(2.7)	2	(11.8)
매우 불만족	0	(0.0)	0	(0.0)	0	(0.0)	0	(0.0)
계	278	(100.0)	115	(100.0)	146	(100.0)	17	(100.0)
평균 (S.D)	3.59	(0.617)	3.54	(0.625)	3.62	(0.600)	3.59	(0.712)

2. 항목별 평가

인구주택총조사 제공 자료에 대한 평가는 대체로 긍정적이다(<표 2-26>). 하지만, 모든 항목에서 ‘대체로 그렇다’라는 응답이거나 ‘보통’수준의 긍정을 보여, 강한 긍정(매우 그렇다)의 비율은 상대적으로 낮았다. ‘매우 그렇다’의 비율이 가장 높았던 항목은 “통계정보를 얻는데 편리함”으로 9.7%였으며, 나머지 다른 항목들은 ‘매우 그렇다’의 비율이 10% 수준에 미치지 못했다.

문항간 평균(5점 척도)을 비교해 보면, “통계정보를 얻는 데에 편리함”에 대한 평가가 3.53점으로 가장 높게 나타났으며, “적절한 시기에 공표됨”(3.36), “적절한 시기에 제공됨”(3.33)이 가장 낮게 평가되었다. 하지만, 문항간 차이의 폭이 그리 크지 않아, 대체적으로 동일한 수준의 평가를 보였다.

<표 2-26> 인구주택총조사 자료 항목별 평가

문항	응답 평균	전혀 그렇지 않음 (1)	대체로 그렇지 않음 (2)	보통 (3)	대체로 그렇다 (4)	매우 그렇다 (5)	합계
		명 (%)	명 (%)	명 (%)	명 (%)	명 (%)	명 (%)
통계정보 얻는데 편리함	3.53	3 (1.1)	25 (9.0)	99 (35.6)	124 (44.6)	27 (9.7)	278 (100.0)
개념/용어 정의 잘됨	3.48	2 (0.7)	22 (8.0)	108 (39.1)	130 (47.1)	14 (5.1)	276 (100.0)
알고자하는 정보를 제공	3.48	5 (1.8)	22 (7.9)	105 (37.8)	127 (45.7)	19 (6.8)	278 (100.0)
이용절차가 편리함	3.43	5 (1.8)	37 (9.7)	107 (38.6)	120 (43.3)	18 (6.5)	277 (100.0)
적절한 시기에 공표됨	3.36	5 (1.8)	26 (9.4)	123 (44.4)	111 (40.1)	12 (4.3)	277 (100.0)
적절한 시기에 제공됨	3.33	3 (1.1)	32 (11.6)	127 (45.8)	101 (36.5)	14 (5.1)	277 (100.0)

3. 소속기관별 평가

인구주택총조사 제공 자료에 대한 평가를 소속기관별로 살펴보면, <표 2-27>과 같다. 중앙부처의 평가가 가장 높은 항목은 “통계정보를 얻는데 편리함”(3.43)이었으며, 다음으로 “알고자 하는 정보를 제공함”(3.37)이 상대적으로 높은 평가를 받았다. 반면, “적절한 시기에 제공됨”이 3.21점으로 가장 낮은 평가를 받았다. 지방자치단체는 “통계정보를 얻는데 편리함”을 3.62점으로 가장 높은 평가를 했으며, 다음으로 “개념/용어의 정의가 잘되어 있음”(3.58), “알고자하는 정보를 제공함”(3.56)을 높게 평가했다. 기타기관은 “개념/용어의 정의가 잘되어 있음”(3.71)을 가장 높게 평가한 것으로 나타났다.

<표 2-27> 소속기관별 인구주택총조사 자료 평가

문항	소속기관				t-값
	전체	중앙부처	지방자치단체	기타기관	
	평균 (S.D)	평균 (S.D)	평균 (S.D)	평균 (S.D)	
통계정보 얻는데 편리함	3.53 (0.831)	3.43 (0.795)	3.62 (0.849)	3.47 (0.874)	1.849
개념/용어의 정의 잘 되어 있음	3.48 (0.746)	3.32 (0.768)	3.58 (0.723)	3.71 (0.588)	2.833**
알고자하는 정보를 제공함	3.48 (0.809)	3.37 (0.729)	3.56 (0.830)	3.53 (1.068)	2.001*
이용절차가 편리함	3.43 (0.825)	3.31 (0.777)	3.51 (0.865)	3.59 (0.712)	1.932
적절한 시기에 공표됨	3.36 (0.784)	3.32 (0.695)	3.39 (0.842)	3.35 (0.862)	0.782
적절한 시기에 제공됨	3.33 (0.787)	3.21 (0.722)	3.41 (0.828)	3.41 (0.795)	2.082*

*: $p < 0.05$, **: $p < 0.01$)

그렇지만, 각 문항에 대해 소속기관별 비교 평가는 큰 의미가 없는 것으로 보인다. 중앙부처의 문항별 평가는 전반적으로 지방자치단체에 비해 낮은 경향성을 띄기는 하지만, 중앙부처와 지방자치단체의 평균차이는 그다지 크게 다르지 않다. 통계적으로 의미 있는 차이를 나타낸 항목은 “개념/용어의 정의가 잘 되어 있음”, “알고자하는 정보를 제공함”, “적절한 시기에 제공됨” 정도이다.

4. 자료 이용강도별 평가

<표 2-28>은 인구주택총조사 자료에 대한 이용강도별 평가표이다. 이용강도가 상대적으로 강한 Heavy User의 경우, “개념/용어의 정의가 잘 되어 있음”(3.57), “알고자하는 정보를 제공함”(3.56), “통계정보를 얻는데 편리함”(3.56)을 높게 평가했다. Light User들은 “통계정보를 얻는데 편리함”(3.48), “알고자 하는 정보를 제공함”(3.38)에 대한 평가가 높게 나타났다.

2) 각 문항의 소속기관별 비교평가에 대한 통계 검증은 사례수가 적은 기타기관(n=17)을 제외하고, 중앙부처와 지방자치단체에 한해 실시하였다.

문항별 Heavy User와 Light User의 평가를 비교해 전반적으로 Heavy User의 평가가 더 좋았던 것으로 나타났다. 이 또한 인구주택총조사를 더 많이 이용할수록 자료이용에 대한 평가도 더 나아지는 현상을 보여주고 있다. 이러한 현상은 자료를 어느 정도 이용해 보지 않으면 접근하기가 어렵기 때문에 나타나는 것으로 보인다. 이런 측면에서, 통계청은 자료 접근이 좀 더 수월하도록 이용편의를 제공함으로써 인구주택총조사에 대한 평가를 높일 수 있을 것이다.

<표 2-28> 인구주택총조사 자료 이용강도별 평가

문항	이용강도		Heavy User		Light User		t-값
	전체		평균	(S.D)	평균	(S.D)	
통계정보 얻는데 편리함	3.53	(0.832)	3.56	(0.914)	3.48	(0.717)	0.870
개념/용어의 정의 잘 되어 있음	3.48	(0.746)	3.57	(0.759)	3.37	(0.718)	2.221*
알고자하는 정보를 제공함	3.48	(0.810)	3.56	(0.848)	3.38	(0.752)	1.807
이용절차가 편리함	3.43	(0.826)	3.54	(0.849)	3.30	(0.779)	2.458*
적절한 시기에 공표됨	3.36	(0.785)	3.43	(0.757)	3.27	(0.813)	1.667
적절한 시기에 제공됨	3.33	(0.789)	3.38	(0.785)	3.26	(0.790)	1.266

주) *: $p < 0.05$, **: $p < 0.01$

제6절. 인구주택총조사에 대한 제안

1. 인구주택총조사에 대한 요구

인구주택총조사에 대한 요구 및 개선 제안을 <표 2-29>에 제시하였다. 먼저, 전체적으로 지방자치단체가 중앙부처보다는 많은 제안을 했다. 제안 내역에는 자료이용의 편의성, 용어 해설 및 문건 제공, 공표 및 제공 시기, 소지역통계나 표본오차 및 집계 정확성 등이 있었다.

이러한 내용을 좀 더 심층적으로 파악하기 위해 연구진은 수원 소재 국책 연구원의 연구원과 서울시 공무원과 약 30분 동안의 면담 형식으로 심층면

접을 하였다. 그들과 나눈 면담에서도 비슷한 내용이 다루어졌으며 이들의 견해는 이 조사와 별로 다르지 않았다.

이들의 제안들은 크게 통계 서비스의 측면과 통계 생산의 측면으로 나누어 볼 수 있다. 자료이용의 편의성, 용어와 문건, 공표 및 제공 시기 등은 서비스에 속하고, 소지역통계나 통계의 정확성은 생산에 속한다. 전체적으로 보면 서비스 측면의 요청이 많았으며, 이들은 오랫동안 지적되어 온 문제들이기도 하다. 또한 이러한 요구는 앞 절에서 살펴본 자료이용에 대한 보통 수준의 평가와도 상응한다.

이러한 결과는 통계청이 통계서비스보다는 통계생산에 더 많은 노력을 기울여왔기 때문에 나타난 것이다. 통계청은 국가통계를 생산에는 상당히 많은 노력을 기울여 왔고 성과도 있었으나 통계서비스에서는 그러하지 못했던 것도 사실이다. 이제 통계의 유용성을 높이기 위해서는 적극적인 통계 서비스 제공 방안을 마련할 필요가 있다. 앞으로, 통계청은 통계생산단계에서부터 통계 유용성과 서비스를 고려하는 방향으로 통계생산체계를 재구축 할 필요가 있다. 이는 요즘 흔히 거론되는 생산자 중심에서 수요자 중심으로의 전환을 의미하기도 한다.

2. 희망 추가조사항목

응답자들이 인구주택총조사에 추가로 조사를 희망하는 항목을 <표 2-30>에 제시하였다. 추가희망 조사항목으로는 개인별 소득, 연령별 교육수준 등이 제시되었다. 이러한 항목들을 모두 수용할 수는 없겠으나, 항상 거론되어온 소득 관련 자료는 일반 이용자들 사이에서도 요구가 많아 언젠가는 조사를 시도해야 할지도 모르는 항목이다. 일부 항목, 예를 들어, 주택 노후화, 사망원인 등에 관한 항목들은 행정자료를 이용하면 어느 정도 보완이 가능하다. 이들 가운데 일부는 현재의 자료로 어느 정도는 계량화할 수 있다. 또 어떤 항목, 예를 들어 결혼이민자의 생활실태나 이혼 후 경제활동과 같은 항목은 새로운 조사나 패널조사가 필요할 지도 모른다.

<표 2-29> 인구주택총조사 개선 요청사항

구분	세부 내용	건의기관
지역세분화	큰 규모 도시에서 동일 시군구내 이동 세분화 필요	서울특별시 도시교통부
정확성/신뢰성	시도별 세대 및 인구분포 등 통계조사의 정확성 높여야	농촌진흥청 안전정책관
	표본조사 - 표본수 늘려서 표본오차 최소화	충북 법무통계담당관
용어 해설	통계용어의 해설 보완	감사원 민자사업팀
공표 시기	조사 시기와 발표 시기의 격차가 커서 사업계획 수립 등에 현실자료로 이용 곤란함	서울특별시기획담당관, 경상북도도시계획과 외 다수
자료 이용	홈페이지 DB 체계 미비, 검색 기능이 미흡, 접근이 용이하지 못함. 자료 구성이 포괄적이어서 세분화된 정보를 찾기 어려움	울산광역시 복지정책과 외 다수
	시스템 개선 - 데이터양이 많을 경우 접속이 불안정함	울산광역시 교통기획과
	Raw Data 이용 편의성 제고	서울특별시 도시관리과
통계 용어	통계 전문 용어, 약자, 세부 통계항목 등에 대한 매뉴얼 제공이 필요	서울시 여성가족정책관 병무청 선병지원팀 외
자료 보완	자료 업데이트 주기 축소	교육인적자원부 기획총괄담당관실
	자료 가공 및 응용이 가능하도록 지원	서울시 균형발전추진본부
	KOSIS DB구성이 시계열(5년단위)로 연결 안됨	사회복지통계과
	상관관계 등 응용통계 자료 확충	광주광역시 민간협력과
	주택, 가구관련 자료 확충	서울시 도시계획상임기획단

<표 2-30> 인구주택총조사 희망 추가 조사항목

구분	세부 내용	건의기관
경제활동	개인별 연평균 소득수준, 연령별 소득수준	강원도 기획관실외 다수
	저소득층 증가율	대구광역시 입법정책담당관
교육	연령별 참여하고 있는 교육형태 및 교육프로그램 관련	교육부
인구	죽보(가계) 등 성씨별 혈연	경제정책국 생활경제과
	혼인연령 및 첫출산 연령, 이혼 및 이혼후 경제활동	서울시 균형발전추진본부
	지역별로 종교분포, 지역별 노인수, 어린이수, 경제활동가능자 수	대구 관광과
	결혼이민자 생활실태 (가구수, 출신국, 결혼기간, 사회적응, 자녀수 등)	전라북도 정책기획관실 여성가족부 외 다수
	사망자 수 및 사망원인	서울시 기획담당관
주택	주택 노후화 (주택 신축후 보유기간)	광주광역시 화재정책팀
	주택형태별 선호도	인천광역시 주택건축과

제7절. 소결

이 조사에서 몇 가지 특징이 나타났다. 먼저 중앙부처와 지방자치단체를 비교하면 지방자치단체가 중앙부처보다 총조사 이용에 조금 더 적극적인 것으로 보인다. 이들의 차이는 표본수가 그다지 많지 않은 편임에도 대체로 통계적으로도 의미 있는 차이를 보였다. 특히 인구주택총조사에 대한 이용경험은 중앙부처의 65.6%에 비해 지방자치단체는 20% 가량 많은 84.4%였다. 이러한 수치는 지방자치단체의 공무원 응답자들은 거의 대부분이 인구주택총

조사 자료를 이용하고 있다는 사실을 보여준다.

이용 빈도를 보면, Heavy User의 비율이 중앙부처에서는 47.0%인데 비해서 지방자치단체에서는 10%p 이상 높은 58.9%였으며, 이 또한 1%의 유의수준에서 통계적으로 의미가 있었다. 이런 측면에서 지방자치단체의 인구주택총조사 이용 수준이 중앙부처보다 높은 것으로 나타났다. 이런 근거에서 인구주택총조사의 이용을 활성화하기 위해서는 먼저 지방자치단체가 필요로 하는 통계생산을 고려할 필요가 있다. 2005년 인구주택총조사에서 시도한 것과 같이 지방자치단체가 필요로 하는 항목을 그들과 지속적인 협력을 토대로 발굴하여 2010년 총조사에 적용하는 것도 한 방안이 될 것이다.

응답자들의 이용에 관한 평가와 제안을 살펴보면 주된 내용은 이용의 편의, 소지역 및 세부 통계, 시의성 있는 통계 등이 있으며, 주된 항목으로는 소득, 혼인, 사망, 외국인에 관한 것들이 있다. 먼저 이용편의에 관해서는 다음 장에서 다룬다. 소지역 및 세부 통계와 시의성 있는 통계는 뒤에 나오는 행정자료 기반의 센서스와 순환조사(rolling census)에서 일부 다루게 된다. 새로운 항목들에 대해서는 다양한 방법으로 수요에 부응할 수 있을 것이다. 그렇지만, 이렇듯 다양한 요구를 모두 충족시킬 수는 없다. 이에 본 연구는 통계청의 통계서비스에 지방자치단체나 중앙부처에서 필요한 통계에 관해 상담할 수 있는 조금은 전문화된 상담창구를 제안한다. 상담창구를 통한 전문화는 통계서비스를 차별화하고, 공공부문 통계서비스의 수준을 질적으로 높이고, 더 나아가 정책에 필요한 통계수요를 파악하는 데 도움을 줄 것이다.

이 조사는 자발적 응답에 의존한 것이어서 응답에서 선택편향(selection bias)이 발생했을 가능성이 있다. 따라서 이 조사의 결과는 일반 공무원들의 생각이기보다는 상대적으로 통계를 많이 이용하는 공무원들의 견해를 반영했을 것이다. 물론 통계수요가 통계를 이용하는 공무원들에게서 주로 발생하기 때문에 이것이 큰 문제가 아닐 것이다. 다만, 심층적 연구의 측면에서 편향성을 통제하는 연구방법은 학문적으로 흥미로운 부분이지만 상당히 심각한 연구가 필요하기 때문에 이 연구에서는 다루지 않은 점이 아쉽다.

마지막으로, 이 조사는 이용자들의 요구를 파악하고, 그에 부응하여 새로

은 센서스 방법들을 모색하기 위한 것이다. 이런 측면에서, 이 연구는 통계 생산자의 입장보다는 수요자 중심에서 인구주택총조사의 미래를 생각한다는 데 최소한의 의의가 있다. 앞으로 인구주택총조사를 포함한 많은 국가통계는 수요자 입장에서 평가하는 통계의 효용가치가 그 통계의 지속적 생산 가능성을 담보해 주게 될 것이기 때문이다.

제3장. 인구주택총조사 활용 증진 방안

제1절. 서론

인구주택총조사의 활용성을 높이는 방안은 통계 생산에서부터 통계 유통 및 서비스에 이르기까지 다양한 부문에 관한 매우 방대한 논의가 될 수 있다. 그러한 방대함은 이 연구의 범위에서 벗어난다. 따라서 여기서는 논의의 범위를 활용성에 관한 직접적인 방안을 다룬다. 특히, 앞 장에서 다룬 이용자들의 개선요구 사항을 중심으로 논의한다.

먼저 인구주택총조사의 활용성을 높이는 방안들을 크게 통계생산부문과 통계 유통 및 서비스 부문으로 나누어 생각해 볼 수 있다. 아무래도 활용과 직접 관계가 있는 부문은 서비스일 것이다. 그러므로 여기서는 통계생산부문에 대해서는 간략하게 언급만 하고 주로 통계서비스 부문을 중심으로 논의하고자 한다. 또한, 앞 장에서 다룬 응답자들의 개선요구 사항도 서비스에 관한 내용이 주를 이루기 때문이다.

먼저 통계생산부문을 간략히 언급하자면, 통계생산은 통계수요를 고려하는 수요자 중심의 생산체계를 구축하는 방향으로 나아가야 한다. 이를 위해 민간기업의 마케팅기법과 같은 방법들을 발굴하고 이를 운용할 인력을 확보해야 할 것이다. 이는 1990년대 초의 우리나라 대기업들과 비슷한 양상이다. 그전까지 대기업들은 기술개발에 역점을 두어 기술이 생산을 주도했지만, 그 무렵부터 마케팅 부문이 생산의 주요 결정을 좌우하기 시작하였다. 이러한 변화는 다양한 취향과 계층으로 성장한 수요자들 변화에 따른 것이며, 그러한 현상이 수요자 중심, 또는 고객 중심의 기업운영으로 이어졌다. 우리나라의 국가통계도 이제 이러한 시점에 다다른 것으로 보인다.

앞 장에서 응답자들의 개선요구에는 통계생산과 관련된 것들이 있다. 정확한 통계나 신뢰성 높은 통계 등 통계품질 관련, 지역을 세분화된 통계, 시의성 있는 통계 등이 포함되어 있다. 정확한 통계나 신뢰성 높은 통계를 생산하기 위해 통계청은 지속적으로 조사방법과 추정방법을 연구·개발해왔다.

더불어, 최근에는 이와 관련하여 국가통계의 품질관리에 많은 노력을 기울이기 시작했다. 또한 최근에 통계청은 소지역통계를 생산하는 방법을 개발하고 있다. 이와 함께, ICT 기술의 활용을 바탕으로 공표시기를 조사 시기에 가깝게 가도록 여러 가지 방법을 사용하여 좋은 성과를 내고 있기도 하다. 이렇듯, 통계생산에서는 다양한 형태의 개선이 있었으며, 앞으로도 이러한 개선의 노력은 지속될 것이다.

한편, 통계서비스에 대해서도 많은 개선이 이루어졌다. 지금 KOSIS 자료를 예로 들면, 시스템의 불편함은 차치하고라도 10여 년 전에는 KOSIS에 접근하는 데 자격제한이 있었던 것이 지금은 인터넷을 통해 누구나 접속할 수 있게 되었다. 하지만, 이러한 개선에도 불구하고 통계이용자들은 통계청 통계 이용에 대해 여러 가지 불만을 토로하고 있다. 설문조사에서 보았듯이 이러한 불만은 주로 서비스 되는 통계의 수준과 관계가 있다. 전에 비해 통계이용자들의 통계 서비스에 대한 요구 수준이 높아진 때문이다.

가장 두드러진 요구는 인구주택총조사 결과를 좀 더 작은 지역을 기준으로 공표해 달라는 것이다. 예전에는 시군구 수준에서도 충분했으나, 이제는 그보다 세분화된 읍면동 또는 그 이하 수준의 통계가 필요해진 것이다. 이러한 요구는 지방자치단체에서 많이 나타나는데, 이는 지방자치단체의 발전에 따라 그들의 정책수립이나 정책연구가 진일보하면서 나타나는 현상이다. 또한, 조사는 되어 있지 않으나 민간기업, 특히 마케팅 분야에서도 그러한 요구가 많다. 이 또한 시장이 세분화하면서 기업들이 시장을 좀 더 상세히 구분(market segmentation)할 필요가 나타났기 때문이다. 이런 측면에서, 여기서는 인구주택총조사 자료의 활성화 방안으로 통계를 공표하는 지역의 범위와 방법에 관해 살펴보고자 한다.

또 하나 두드러지게 높아진 통계수요는 원자료(microdata)에 관한 것이다. 이 또한 우리나라의 연구역량이 커짐에 따라 연구자가 자신에게 필요한 정보를 가공해서 사용할 필요성이 증가한 때문에 발생한다. 이러한 수요는 단지 연구원이나 학자들뿐만 아니라 지방자치단체와 중앙부처에서도 발생하였다(<표 2-19> 참조). 이는 우리의 행정에서도 정책적 목적에 따라 자료를 직접 가공해서 사용할 수 있는 인력이 증가했을 뿐만 아니라 그리해야 할

필요성이 증가한 때문이다. 이런 측면에서, 여기서는 원자료 제공을 중심으로 살펴보고자 한다.

제2절. 외국의 통계 집계 및 공표구역³⁾

1. 전반적인 경향과 특징

외국에서도 사회발전에 따라 다양한 형태로 소지역통계에 대한 수요가 증가함에 따라 통계 생산과 공표를 위해 공간적으로 적절한 규모를 정하여 전국을 여러 형태의 구역으로 나누어 놓았다. 이러한 구역화는 크게 조사를 위한 조사구역과 집계 및 공표를 위한 공표구역으로 나눌 수 있다. 이외에도 이러한 구역들의 기초 단위가 되는 최소의 구역을 정해 놓기도 한다. 여기서는 외국에서 정해 놓은 통계 관련 지역들을 공표구역 중심으로 살펴본다.

<표 3-1>은 미국, 영국, 캐나다, 호주, 일본 5개국에서 사용하고 있는 통계구역을 요약한 것이다. 전체적인 모습을 요약하면, 공표구역의 경우 일본이나 우리나라와 같이 기존의 행정구역 체계를 그대로 사용하는 경우도 있으나, 대부분의 국가에서는 행정구역과 별도로 통계의 집계 및 공표를 위해 새로운 지역 구분 체계를 사용하고 있다. 각국의 통계구역 체계는 국가에 따라 다양하게 나타나고 있기 때문에 하나로 단일하게 설명할 수는 없지만, 몇 가지 공통된 특징이 나타나고 있다.

첫째, 조사구역과 공표구역의 분리 경향이다. 대부분의 국가들, 그 가운데도 특히 영연방 국가들은 과거에 조사구역을 공표단위로도 사용하였다. 그러나 최근에는 조사구역과 공표구역을 분리하여 사용하고 있다. 전국을 작은 구역으로 나누어 관리하는 것은 매우 복잡하고 지난한 일이었지만, 최근에 Geographical Information System의 발달로 이러한 어려움이 많이 줄었기 때문에 다양한 형태의 구역을 지정하여 관리할 수 있게 되었다. 조사구역은 통계 조사의 효율성을 목적으로 하며, 공표구역은 공표를 목적으로 한다. 공표구역의 경우 개인정보보호, 통계적 유의성, 소지역 통계에 대한 수요 등을

3) 강영욱·윤은주·원종석·정재희(2007)을 바탕으로 재구성

다양한 요건을 만족시키는 기준을 마련하고 있다.

둘째, 공표 구역의 설정 조건으로는 인구 기준이 가장 중요한 요소로 작용하고 있으며, 기존의 행정구역 체계 및 기타 지역 구분 체계와의 통합가능성을 고려하고 있었다. 미국의 경우에는 블록그룹과 센서스트랙을 인구 조건만을 기준으로 설정하였으며, 영국과 캐나다의 경우는 인구 조건을 기본 조건으로 형상 조건, 내부적 동질성, 도로에의 접근성 등을 추가적으로 고려하였다.

셋째, 일반 행정, 선거구 획정, 도시지역 구분 등 에서도 원활히 사용할 수 있도록 공표구역들을 행정구역 경계를 넘나들지 않도록 작성하고 있다. 통계 기반 구역의 용도를 살펴보면 미국은 최소 공표 단위인 블록을 도시화 지역, 선거구, 학군 등 다양한 지역 설정의 기본단위로 활용하고 있다. 호주는 Collection District 설정결과를 바탕으로 이들 여럿을 하나로 묶어 도시화 지역을 설정한다. 캐나다에서는 대도시권인 Census Metropolitan Areas의 설정에서 Census Subdivision을 기본단위로 하여 통계공표지역인 Dissemination Area와 연계하여 사용하고 있다.

넷째, 소지역 통계 수요의 증가에 따라 자료의 유형에 따라 공표기준을 다르게 적용하여 다양한 형태로 통계를 공표하고 있다. 특히, 개인정보노출 위험이 적은 자료는 통계적 의미 있는 최소단위까지 자료를 제공하여 소지역 통계의 활용도를 높이고 있다. 이와 관련하여, 자료 공표 시 우려되는 개인정보보호를 위한 제도적 장치를 도입하고 있다. 예를 들어, 미국과 일본은 표본 조사의 공표 단위를 전수 조사의 공표 단위보다 상위의 지역 단위에서만 발표하고 있으며, 영국에서는 교차표분석 자료의 공표 시 인구 1,000명을 기준으로 하여 최소공표단위인 Output Area 수준에서는 발표를 금지하고 있다. 호주는 공표구역설정에서 최소인구기준을 100명, 캐나다는 400명으로 정하여 구역 획정에서부터 개인정보를 보호할 수 있도록 하고 있다.

<표 3-1> 외국의 통계 관련 구역의 특징*

	미 국	영 국	캐나다	호 주	일 본
센서스의 시행	10년 단위 (○○○○년)	10년 단위 (○○○1년)	5년 단위 (○○○1, ○○○6년)	5년 단위 (○○○1, ○○○6년)	5년 단위 (○○○○, ○○○5년)
최소조사 구역 및 도입년도	Block/ 1940년	Enumeration District/ 1841년	Enumeration Area/ 1961년	Census Collection District	조사구 (Enumeration District)
조사구역 설정기준	① 최소면적 약 2,780ha ② 최소폭 약21m ③ 최소1면 이상 도로에 인접	① 조사원의 업무량 표준화 ② 주요간선도로, 강, 철도 등으로 구분 ③ ward, parish의 경계 유지	① 조사원의 업무량 표준화 ② 단순한 형상 ③ 도로에 의한 접근성 ④ 타 공간단위의 경계	① 조사원1명의 10일간 업무량 ② SLA및 타지역 구분의 경계 유지 ③ 가시적 지형지물로 구분	① 조사원의 업무량 기준 ② 기본단위구를 결합하여 형성
조사구역 인구규모	제한 없음	지역별로 다양 (평균 450명)	농촌지역 125명 도시지역 650명	최소 100명 원주민 지역 80명	평균 50명
최소공표 구역 및 도입년도	Block/1940년 (주공표는 Census tract)	Output Area/ 2001년	Dissemination Area/ 2001년	Census Collection District (2011년부터Mesh Block)	cho, aza Grid Square/ 1973년
공표구역 설정기준	조사구역과 동일	①인구:최소 100명,최적 312명 ②형상 : 중심간 거리의 제곱을 최소화 ③내부 동질성: 주택유형 및 점유형태 ④ 타 공간단위의 경계	① 인구:최적500명 ② 형상:정방형우선 ③ 타공간단위의 경계 ④ 가시적 경계의 사용여부	조사구역과 동일	①국세조사: 행정구역 체계 ②Grid Square:위도 40' 경도 1°에 해당하는 격자를 80×80으로 균등 분할
가시적 경계의 사용여부	행정구역 및 도로, 철도, 하천, 능선 등의 가시적 경계 사용	행정구역 경계를 사용	행정구역 및 도로와 같은 가시적 경계를 사용	행정구역 및 도로, 하천 등의 가시적 경계 사용. 농촌의 경우는 도로를 경계로 사용하지 않음	행정구역 경계를 사용
공표구역 인구규모	인구 기준 없음	평균 297명	400~700명	도시지역 평균 512명	해당사항 없음
개인정보 보호를 위한 규정	표본조사는 블록그룹 (600~3,000명)단위로 공표 일부 교차표분석 자료는센서스트랙 (1,500~8,000명) 단위로 공표	교차분석 자료는 1,000명 이상의 지역에 대해서만 발표	최소 공표단위의 인구 최소기준을 설정(400명 이상)	최소 공표단위의 인구 최소기준을 설정(100명 이상)	표본조사는 cho, aza의 상위체계인 시구정촌 단위로 공표
조사구과 공표구역의 관계	동일	분리. ED를 통해 수집된 자료를 개별주소 단위로 입력	분리. EA와 DA의 기본단위인 Block을 통해 연계가능	동일	조사구역이 공표구역의 하위
도시/비도시 구분	블록을 도시화지역 설정의 기본단위로 사용	도시지역의인구측정에 OA인구를 이용	구분. 도시지역은 DA의상위체계인 CSD 단위로 설정	CD의 설정결과를 기준으로 도시화지역을 설정	구분하지 않음 행정구역에 따른 구분은 가능

출처 : 강영옥 · 윤은주 · 원종석 · 정재희(2007)

2. 미국의 통계구역 설정 현황

가. 미국의 센서스와 통계구역 구분체계

미국은 매 10년마다 연도 끝이 '0'으로 끝나는 해에 센서스를 시행한다. (그림 3-1)은 미국의 통계 관련 구역을 위계적 관계로 표시한 것이다. 미국의 통계구역(statistical area) 체계는 센서스 트랙(Census Tracts)을 중심으로 상위에는 카운티(County), 주(State) 등의 행정 단위로, 하위에는 블록그룹(Block Group), 블록(Census Block) 등으로 체계화 되어 있다. 센서스 트랙은 행정 단위인 카운티와 주의 지역 경계 내에서 설정되며, 센서스 트랙의 하위단위인 블록그룹과 블록은 센서스 트랙을 분할하여 구획한 것이다. 통계 자료는 센서스 트랙, 블록그룹, 블록 등 다양한 스케일의 통계구역별로 공표된다.

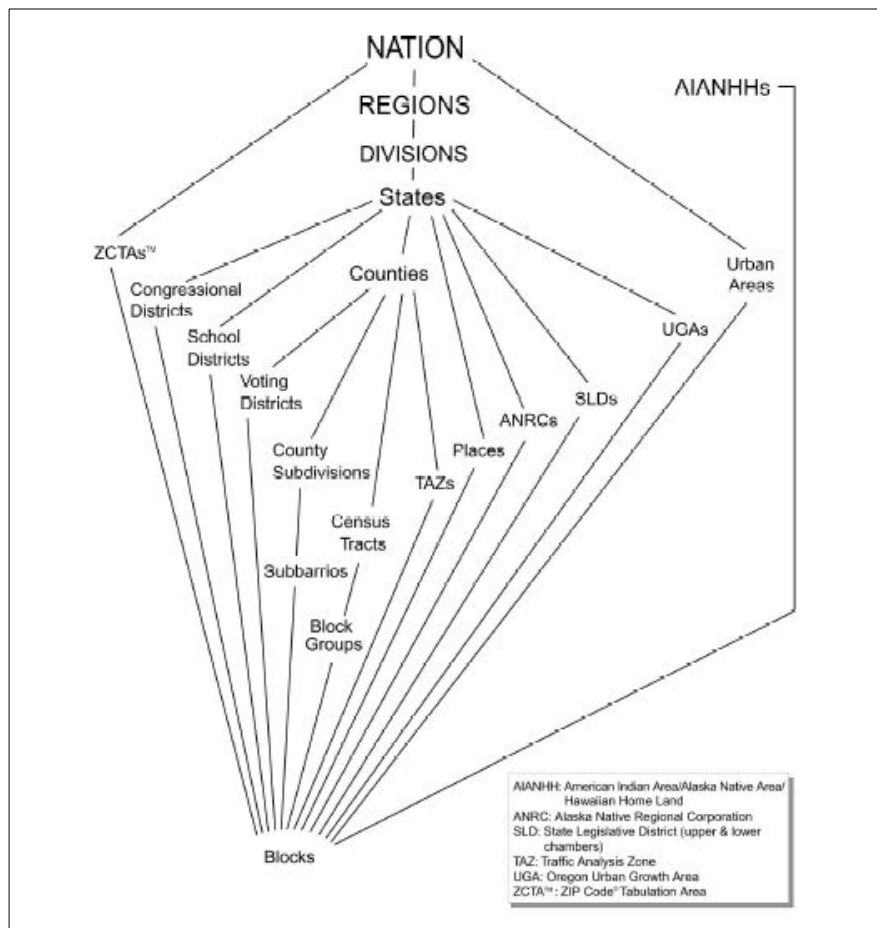
센서스 자료의 공표 체계는, 전수조사의 경우, 총 286개 테이블 중 171개 인구 항목 테이블(P항목)과 56개 주택 항목 테이블(H항목)이 블록 단위까지 공표되며, 59개의 인구 항목 테이블(PCT 항목)은 센서스 트랙 단위까지 공표된다. 표본조사의 경우 총 813개의 테이블 중에서 인구항목 196개(PCT), 주택항목 208개(HCT)가 센서스 트랙 단위까지 공표되며, 나머지 409개 테이블 자료는 블록그룹 단위까지 공표된다⁴⁾. 전수조사와 표본조사의 공표내용은 총 1,099개의 테이블 자료로서, 기본 항목뿐만 아니라 여러 항목들간의 교차분석(Cross-tabulation) 형태로 제공되기도 한다. 센서스의 최소 조사단위로 는 통계구역 체계의 가장 최소단위인 블록이 이용된다.

나. 센서스 트랙(Census Tracts)

센서스 트랙은 카운티(county) 내에 위치한 준영구적인 소지역을 뜻하며, 준영구적인 지리적 공간에 대해 유의한 통계의 생산과 센서스 간 데이터 비

4) 센서스 트랙 단위까지만 공표되는 데이터는 주로 인종·민족, 거주 기간에 따른 구분 집단 내에서의 교차분석 항목과 가계 소득과 관련된 교차분석 항목 등이다.

교를 가능케 하는 소지역 통계구역으로서의 기능을 담당한다. 센서스 트랙은 Walter Laidlaw에 의해 도입된 개념으로 1910년 센서스에서 최초 인구 5만 이상의 8개 도시에 적용한 이후 1930년 센서스에서 18개 도시로 확대되었으며, 1940년부터 센서스의 표준 공간단위로 채택되었다. 1990년 TIGER (Topologically Integrated Geographic Encoding and Referencing) DB가 전국 단위로 구축되면서, 도시 지역은 센서스 트랙으로, 비도시 지역은 블록 넘버링 구역(Block Numbering Area, BNA)으로 설정하였으나, 2000년 센서스에서는 센서스 트랙으로 단일하게 통합하였다.



(그림 3-1) 미국의 통계구역 구분 체계

출처 : 미국 통계국 <http://www.census.gov>

다. 블록그룹과 센서스 블록

(1) 블록그룹(Block Group)

블록그룹은 전수조사와 표본조사를 위한 공간적 단위로, 1970년 센서스부터 부분적으로 도입되었으며, 1990년 센서스 때 전국 범위로 설정되었다. 1990년 센서스에서는 가구 수를 기준으로 블록그룹을 구분하였으나, 2000년 센서스부터는 인구를 기준으로 설정하고 있다.

블록그룹은 센서스 트랙을 분할하여 만들어지며, 하나의 센서스 트랙은 최대 9개의 블록그룹으로 나뉜다. 블록그룹은 센서스 트랙을 분할하여 만들어지기 때문에 센서스 트랙의 경계는 블록그룹의 경계가 되며, 블록그룹의 경계 속성 조건은 센서스 트랙의 경우와 동일하다. 또한 블록그룹은 센서스 트랙과 마찬가지로 인구와 경계 조건을 만족하여야 하며, 단순한 형태, 도로에 의해 모든 지점에 접근이 가능한 인접한 지역이어야 한다.

<표 3-2> 지역별 블록그룹의 인구 기준

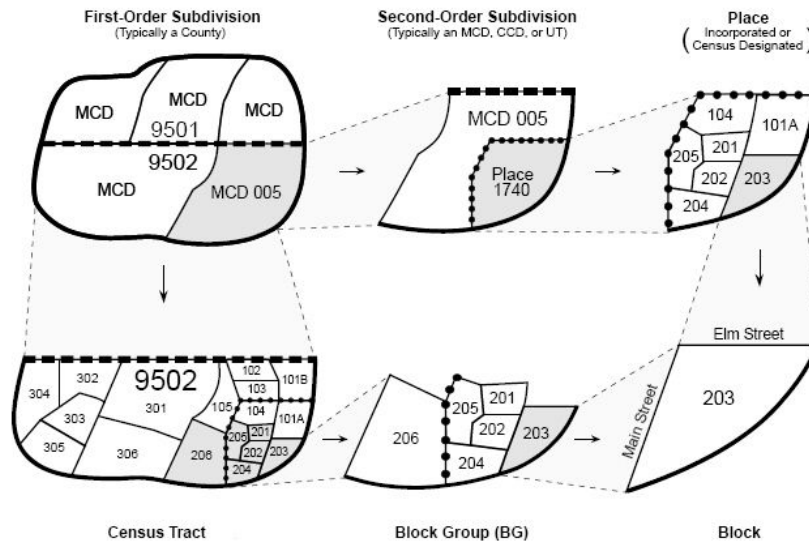
지 역	기준(명/가구)		
	최 적	최 소	최 대
미국 본토, 푸에르토리코, 버진아일랜드, 사모아, 괌, 북마리아나제도	1,500 (600)	600 (240)	3,000 (1,200)
인디언보호구역	1,000 (400)	300 (120)	3,000 (1,200)
특별시설(special place)	없음	300	1,500

출처 : U.S. Department of Commerce et al.(2000), 강영옥외(2007)에서 재인용

(2) 센서스 블록(Census Block)

센서스 블록은 통계자료의 조사와 공표의 최소단위로 1940년 센서스에 처음 도입되었으며, 1990년 센서스에서 전국으로 확대되었다. 블록의 구획은 TIGER DB를 기반으로 인구 규모에 관계없이 나눌 수 있는 가장 작은 크기

로 구획되기 때문에 인구가 전혀 없는 블록도 발생한다. 1990년 센서스의 경우 블록 수는 총 7,017,427개였으며, 이 중 약 200만 개의 블록(전체 블록의 약 28%)이 인구가 전혀 없었다.



(그림 3-2) 미국의 소지역 통계구역의 구분과 번호체계 출처 : 강영옥외(2007)에서 재인용

3. 영국의 통계구역 설정 현황

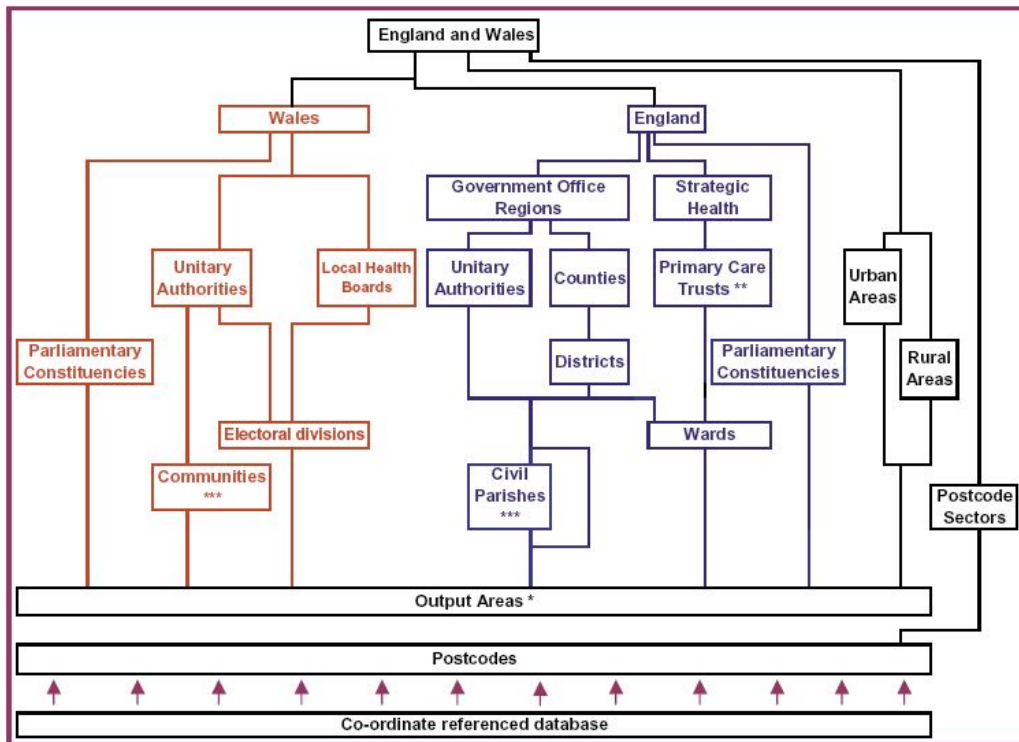
가. 영국의 센서스와 통계구역 구분체계

영국의 센서스는 10년 단위로 이루어지며, 가장 최근의 센서스는 2001년에 실시되었다. 영국의 센서스 및 통계구역 구분 등은 잉글랜드와 웨일즈 지역이 같은 체계로 이루어져 있고, 스코틀랜드와 북아일랜드 지역은 독립적으로 통계 구역의 설정과 집계, 공표가 이루어지고 있다⁵⁾.

통계구역 체계는 기존에 설정되어 있던 법적·행정적 경계인 ward, parish, community 등이 그대로 사용되며, 추가적으로 통계적 목적을 위한 소지역 단위로 Output Area(OA)와 Enumeration District(ED)가 설정되어

5) 이 글에서 특별한 언급이 없으면, 영국은 잉글랜드와 웨일즈 지역을 가리킨다.

있다. OA와 ED는 상위에 속하는 ward, parish, community와 같은 행정 단위의 경계를 따르기 때문에 OA와 ED 규모에서의 통계자료는 손쉽게 더 큰 지역단위로 변환해서 이용할 수 있다.



(그림 3-3) 영국의 통계구역 구분 체계

영국의 센서스는 2001년 센서스부터 통계 자료의 조사 단위와 공표 단위가 분리되어 있는데, 조사 단위로는 ED가 사용되며, 공표 단위로는 OA가 사용된다. ED와 OA는 모두 ward, parish, community 등의 경계를 따르지만, 둘 간의 경계는 연관성이 없다. ED는 조사원의 조사구역으로서, ED를 통해 수집된 센서스 자료는 개별 주소 단위로 입력되며, 이 자료가 OA를 기본 단위로 집계, 공표된다.

센서스 자료의 공표는 2001년에 처음 도입된 OA를 기본단위로 하지만 개인 정보의 보호를 위해 표준 테이블 자료(교차표분석이 이루어진 센서스 자료)는 ward를 기본단위로 공표된다. 영국에서는 센서스 자료의 공표와 관련

하여 최소공표인구를 설정하였는데, 2001년 센서스의 경우, 일반 센서스 자료는 인구 100명이고 40가구 이상일 경우, 표준 테이블 자료는 인구 1,000명 이상이고 400가구 이상일 경우에만 자료를 공표하게 규정하였다. 예를 들어, 최소 인구가 100명, 평균인구가 약 300명인 OA에서는 표준 테이블 자료는 발표되지 않는다.⁶⁾

나. Enumeration District

ED는 OA의 도입 이전인 1990년 센서스까지 통계 자료의 조사·집계·공표의 기본 단위로 사용되었으나, 현재는 통계 자료의 조사 단위로만 사용되고 있다. EA는 1841년에 조사 단위로 처음 도입되었으며, 1961년 센서스부터 공표 단위로도 사용되었다. 1991년 센서스에서는 잉글랜드, 웨일즈 지역에 총116,919개의 ED가 설정되었으며, 추가적으로 4,840개의 '특수' ED(교도소, 요양원 등)가 설정되었다.

ED의 경계는 ED가 원래 통계 자료의 조사 단위로 도입되었기 때문에 조사원의 업무량을 균등하게 하면서, 통계조사를 가장 효율적으로 배포하고 수집할 수 있도록 설정되었다. 따라서 주택 간의 간격이 먼 시골 지역의 경우, 조사원의 업무량을 고려해 도시 지역에 비해 더 작은 가구 수 단위로 ED가 설정되었다. ED는 주요 간선도로, 강, 철도 등을 경계로 삼으며, 공표단위로서 상위의 행정단위로의 집계가 용이하도록 ward, parish, community의 경계를 따른다. ED의 경계는 이와 같은 원칙에 따라 담당관이 대축척 지도에 수작업으로 구획을 하게 되며, 1991년 센서스부터는 GIS 기술의 발전에 따라 ED의 경계를 디지털화 하였고, 2001년 센서스부터는 수작업 대신 주소 자료와 연계된 GIS 기반의 디지털지도로 구획이 이루어졌다.

ED는 본래 통계 자료의 공표를 목적으로 구획된 것이 아니었기 때문에 공표 단위로 사용하기에는 여러 가지 문제점을 가지고 있었다. 먼저, ED는 통계 자료의 조사 단위로서 조사자의 조사가능 범위를 기준으로 구획하였기 때문에 형상과 인구 규모가 매우 다양하게 나타난다. 또한 경계 설정 시 ED 내부의 사회적 동일

6) 영국 통계청, <http://www.statistics.gov.uk/census2001>

성 등은 고려되지 않기 때문에 각종 연구에 통계자료를 활용함에 있어 어려움이 있었다. 다음으로, 센서스 간의 ED 설정에 있어서 경계의 변화를 최소화하는 고려는 있었지만, 업무량의 변화에 따라 ED의 구획이 센서스에 따라 크게 달라지는 문제가 발생하였다. 한 예로 1981년 센서스에서 설정되었던 ED 중 68%가 1991년 센서스에서 새롭게 설정되어 두 센서스 간의 시계열적 비교 분석이 어려웠다. 또한, 1991년 센서스의 경우 공표 가능한 최소인구와 가구수의 기준으로 설정되었던 50명, 16가구에 미달하는 ED가 다수 존재하여, 데이터의 이용에 많은 제한이 있었다⁷⁾. 이 외에도 센서스 이외의 통계에서 사용 빈도가 높은 우편번호체계와 연동이 되지 않는다는 점 역시 통계 자료의 공표 단위로서 ED 사용상의 문제점으로 들 수 있다. 따라서 이러한 문제점들을 보완하고 해결하기 위해 조사 단위와 공표 단위 간의 구분에 대한 요구가 커졌으며, 이에 새로운 공표구역으로 2001년 센서스에서 OA가 도입되었다(Martin, 2002).

다. Output Area

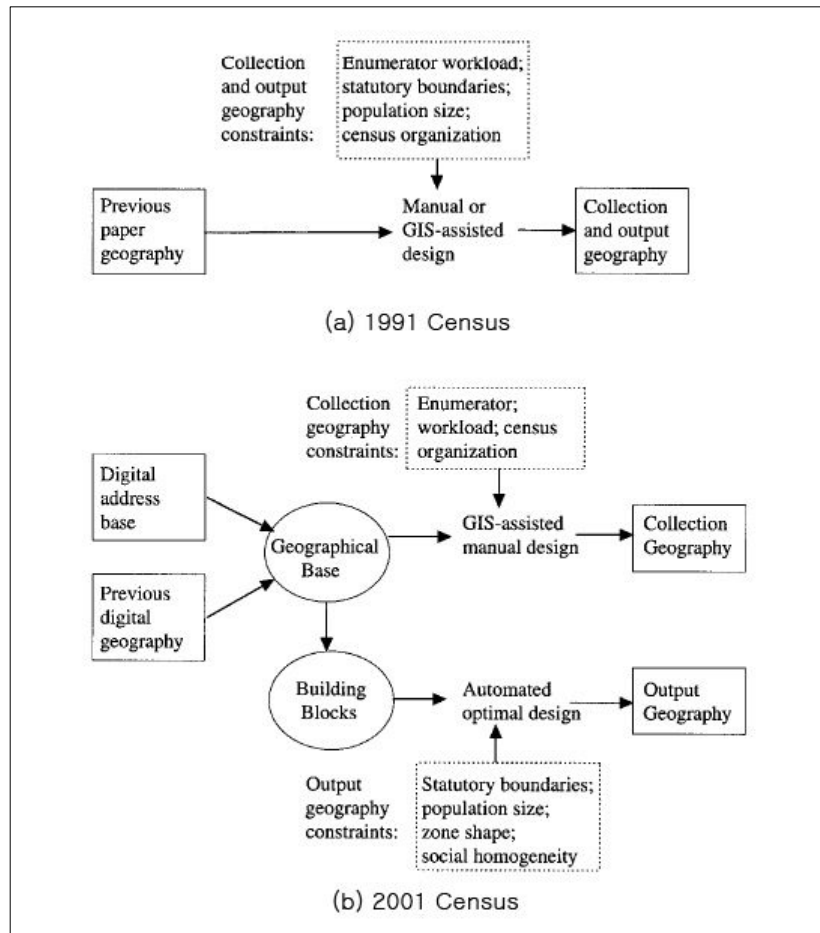
OA는 2001년 센서스부터 도입된 통계의 공표 단위로, 기존에 공표단위로 사용되던 ED의 기능을 대체하였다. 하지만 OA는 ED의 기존 기능 중 통계의 공표 단위로서의 기능만 대체하였으며, 통계 자료의 조사 단위로는 ED가 계속 사용되고 있다.

OA의 도입 목적은 1991년 센서스에서 나타났던 최저 인구 기준 미달과 같은 문제를 해결하고, 우편번호 지역과 센서스 지역의 연동이 가능하게 하기 위함이며, 공표 구역의 설정에 있어 지역 단위의 인구·사회적 동질성·형상 등을 표준화하기 위한 목적도 포함하고 있다(Martin, 2002).

OA는 지방정부 관할구역 내 근린 수준에서 센서스 및 기타 통계를 공표하는 공간단위로서 활용됨으로써 소지역 수준에서 중앙정부 및 지방정부가 집행해야 할 정책의 대상과 그 공간범위의 설정에 필요한 사회·경제적 데

7) 이러한 소규모 ED의 경우 '제한' ED로 정의하여 인구 및 가구수를 제외한 모든 값에 no data를 부여하고, 나머지 통계는 인접한 ED에 병합하였다. 1991년 기준으로 542개의 ED가 50명의 최소 인구기준에 미달하여 인접한 ED에 통합되었으며, 2001년 센서스의 경우에는 최소기준이 100명, 40가구로 상향되어 기준에 미달한 ED가 1,400개에 달하였다.

이터를 제공할 수 있다. OA가 도입됨으로써 정책 대상 및 그 공간범위의 설정을 정교하게 수행할 수 있어 정책 집행의 효율성을 높일 수 있으며, OA 별로 공표되는 센서스 데이터를 토대로, 정책의 효과에 대해 소지역 수준에서 모니터링을 수행하는 것도 가능해진다. 또한 OA는 통계자료의 제공 뿐 아니라 선거구 획정 등과 다른 지역 구분에 필요한 소지역 센서스 자료를 제공할 수 있다(변필성외, 2006).



(그림 3-4) 조사구역과 공표구역의 분리
출처 : Martin(1998b), 강영옥외(2007)에서 재인용

2001년 센서스의 경우 설정된 OA 개수는 총 175,434개이며, OA당 평균인구는 297명, 평균 가구 수는 124가구였다. OA의 번호는 ward 내에서 고유한 네 자리의 번호를 순서대로 부여되는데 county, district, ward의 번호를 포

함하여 8자리의 코드로 구성된다. 즉, 00-AA-BB-0001 같은 형식이다.

라. Super Output Area

SOA는 근린통계를 위한 안정적이고 통일된 단위로서, OA를 기초로 만들어진 공간 단위이다. SOA의 도입 이전까지 근린 통계(Neighbourhood Statistics, NeSS)의 공간 단위로는 선거구가 사용되었는데, 선거구의 경우 지역에 따라 인구의 편차가 100명에서 30,000명 이상까지 매우 크게 나 국가 전체 차원에서 지역 간 비교에는 불합리한 측면이 존재하였다. 또한, 선거구의 경우 잦은 경계의 변화가 발생하여 시계열 분석이 어려웠으며, 특히 인구가 적은 지역의 경우 개인정보 보호를 위해 통계자료를 발표하지 못하는 경우도 많았다. 이러한 문제의 해결을 위해 일관성 있고 영구적인 공간 단위의 도입이 요구되었으며, 그 일환으로 2001년 센서스에서 OA가 도입되었다(Martin, 2004).

SOA는 공간 규모면에서 OA와 지방 정부 행정경계의 중간 정도에 해당하는 크기이며, lower layer, middle layer, upper layer의 총 3개 층으로 이루어진다. 현재는 lower layer와 middle layer까지 설정이 완료되었으며, upper layer는 2006년 중에 설정될 예정이다. SOA는 2004년부터 2006년까지 2년간의 테스트 기간을 통해 경계 설정의 타당성, 인구 변화에 따른 경계 조정 방법 등에 대한 검토가 이루어진 이후에 최종 확정될 예정이다⁸⁾.

Lower layer는 최소 인구 1,000명, 평균 인구 1,500명으로, 4-6개 정도의 OA를 묶어서 설정한다. 2001년 센서스의 Standard Table(ST) wards의 경계를 준수하며, 인구 규모, 상호 근접성 및 동질성을 바탕으로 컴퓨터를 통해 자동으로 설정한다. 2004년 2월에 설정되었으며, 총 34,378개가 구축되었다. Middle layer는 최소 인구 5,000명, 평균인구 7,200명으로 몇 개의 lower layer를 묶어서 설정한다. 2003년 기준의 지방 정부 경계를 준수하며, lower

8) SOA와 비슷한 공간 단위로 스코틀랜드와 북아일랜드에는 다른 이름의 지역구분이 존재한다. 스코틀랜드의 경우 인구 500-1,000명 규모의 Data zone과 2,500-6,000명 규모의 intermediate zone이 있으며, 북아일랜드에는 인구규모 1,300-2,800명의 Super Output Area가 있다.

layer와 같은 방식으로 초안 작성 후, 지방 정부 및 기관의 검토를 거쳐 확정된다. 2004년 8월에 설정되었으며, 총 7,193개가 구축되었다. Upper layer는 최소 인구 25,000명 선으로 계획되고 있으며, 세부사항은 아직 확정되지 않았다.

4. 캐나다의 통계구역 설정 현황

가. 캐나다의 센서스와 통계구역 구분체계

캐나다의 센서스는 5년 간격으로 이루어지면, 가장 최근에는 2001년에 이루어졌다. 캐나다의 지역구분체계는 크게 SGC(Standard Geographical Classification)와 SAC(Statistical Area Classification)로 나눈다. SGC는 센서스를 비롯한 기본적인 통계의 조사와 공표를 위해 구분된 체계이며, SAC는 도시지역 구분과 그에 따른 통계자료의 활용에 이용되는 체계이다.

캐나다의 센서스 데이터는 SGC 체계를 기본으로 발표되는데, 통계 자료의 최소 공표단위로는 2001년 센서스부터 도입된 Dissemination Area가 사용된다. 대부분의 통계 자료는 DA를 기준으로 공표가 이루어지며, 일부 데이터의 경우, 더 하위의 블록 단위까지 발표되기도 한다. 센서스의 조사단위로는 Enumeration Area가 사용된다.

SGC는 크게 주/테리토리(province/territory) - CD(census division) - CSD(census subdivision)의 위계로 이루어져 있으며, 전 국토에 걸쳐 설정되어 있다. 주와 테리토리는 총 13개이며, CD는 288개, CSD는 5,600여개가 설정되어 있다. CD와 CSD는 새롭게 설정된 지역 구분이라기보다는 법적, 행정적으로 존재하는 비슷한 규모의 수많은 지역 단위를 단일하게 표현하기 위해 도입한 개념으로, CSD는 기초자치단체(municipality)와 그에 준하는 모든 지역을, CD는 주와 기초자치단체의 중간 규모에 해당하는 모든 지역을 가리킨다.

구체적으로, CD는 division, county, district, district municipality, Metropolitan municipality, Region, Regional district, Regional municipality,

United county, Communauté urbaine, Municipalité régionale de comté 등 11개 행정 단위를 나타내는 지역으로, 이들은 주 법률에 의해 설정되거나, 주 법률에서 정하지 않는 지역(Newfoundland and Labrador, Manitoba, Saskatchewan, Alberta, Yukon Territory, Northwest territory, Nunavat)은 통계청에서 지역 상황에 맞게 설정한다. CSD는 city, town, village, Indian reserve, Indian settlement, unorganized territory 등 기초자치단체에 준하는 46종류의 행정 단위를 나타내는 용어로, CSD의 종류는 CSD의 이름 뒤에 병기함으로써 CSD를 구분한다. 이 외에 인디언 보호구역, 인디언 거주지 등도 개별적으로 CSD를 형성하는데, 약 2,800개 중 1,052개의 인디언 보호구역과 28개의 인디언 거주지가 개별 CSD로 지정되어 있다.

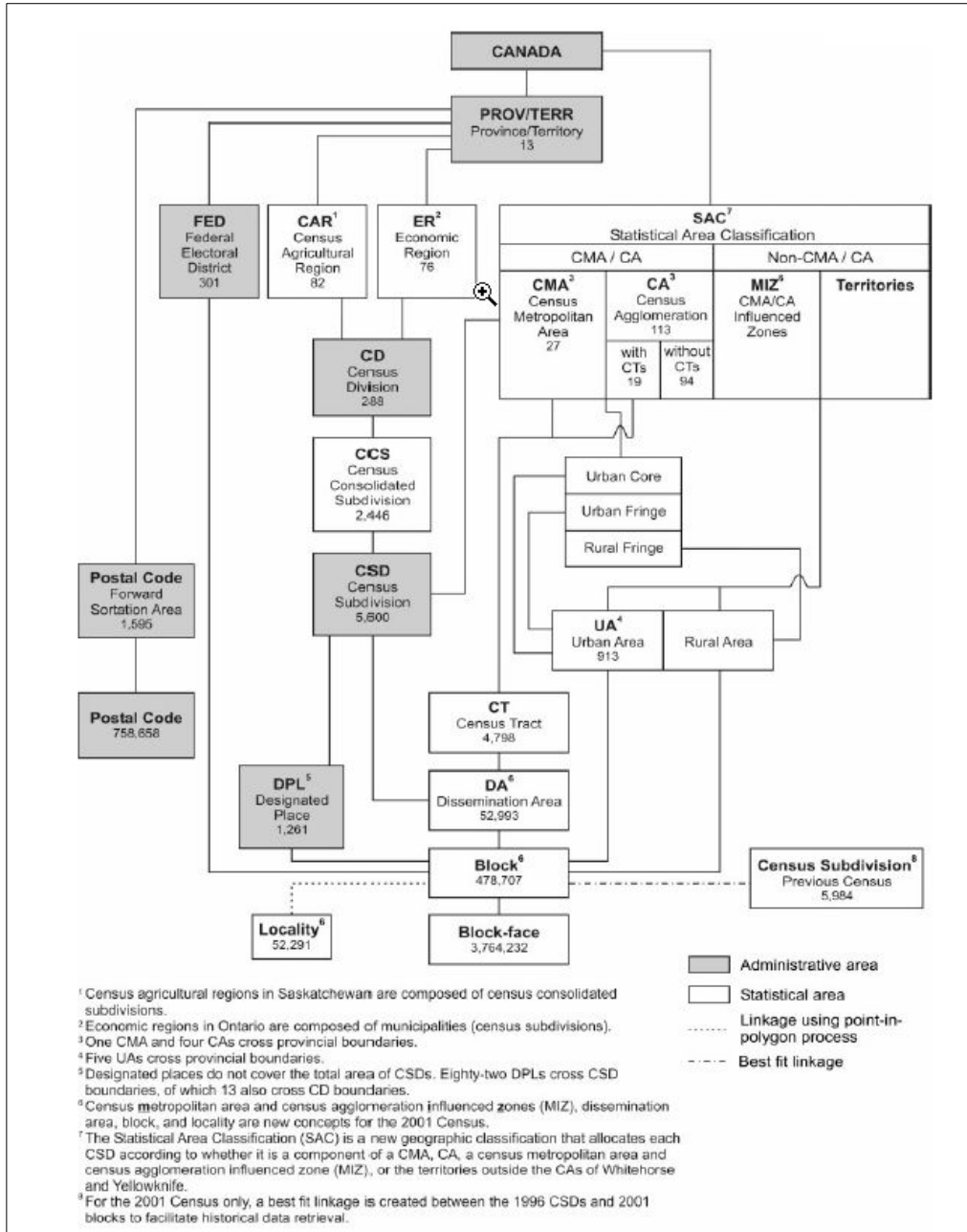
나. Enumeration Area

EA는 센서스를 위해 확정된 공간 단위로 인구 규모는 125명(농촌 지역)~650명(도시 지역) 정도이다. 영국의 ED와 마찬가지로 1996년 센서스까지는 조사단위 뿐 아니라 최소 공표단위로도 사용되었으나, 2001년 센서스부터 DA가 도입되면서 EA는 조사단위로만 사용되고 있다.

EA의 구획은 National Geographic Base에 기반 해서 형상의 단순화, 도로에 의한 접근성, 다른 공간 단위의 경계, 인구 조건 등을 고려하여 자동으로 이루어지는데, 도시지역(urban area)과 DA를 제외한 모든 지역 단위의 경계를 준수하게 된다. 대형 아파트의 경우 하나의 건물이 여러 개의 EA로 분할되기도 하며, 선박의 경우 모항의 EA에 포함된다.

EA는 영국의 ED와 마찬가지로 최초 목적이 통계자료 조사의 편의성을 위한 것이었기 때문에 공표단위로 사용하는 데에는 여러 가지 문제점이 있었다. 대표적인 예로, 1996년 센서스의 경우 개인정보 보호를 위해 250명 이하의 EA의 소득에 대한 데이터와 40명 이하의 EA의 모든 데이터의 공표를 금지하였는데, 전체 EA의 27%가 250명 이하, 10%가 40명 이하라는 조건에 해당되어 통계자료의 공표가 이루어지지 못하는 문제가 발생하였다. 따라서 조사단위와 분리된 새로운 공표단위의 필요성이 제기되었으며, 2001년 센서

스부터는 DA가 새로운 공표단위로 도입되었다(Statistics Canada, 2003).



(그림 3-5) 캐나다 통계구역(Statistical Area) 체계

출처 : Statistics Canada(2004: 205)

다. Dissemination Area

DA는 2001년 처음 도입된 통계 자료 공표의 가장 최소 단위로, 하나 혹은 그 이상의 블록을 결합하여 설정한다. DA는 비교적 변화가 없는 안정된 공간 단위로 인구규모는 400~700명이며 최적인구는 500명이다. 하지만 2001년 센서스의 경우 여러 가지 실행상의 문제로 인해 larger urban centre⁹⁾ 이외 지역은 2001년 센서스의 EA의 경계를 그대로 사용하였으며, CSD와 센서스 트랙의 경계를 준수하기 위해, 400명 미만의 인구지역도 존재하였다.

DA의 구획 방식은 영국의 공표단위인 OA의 구획 방식과 비슷한 과정을 통해 이루어지는데, 여기에는 GArDS(Generalized Area Delineation System)이라는 방법이 사용된다. GArDS는 블록을 빌딩블록으로 하여, DA를 구획하는 최적의 조건들을 설정하고 컴퓨터 알고리즘을 통해 각 블록을 다양한 방식으로 조합한 후, 설정된 최적 조건에 가장 가까운 블록의 조합을 선택하는 방법이다. 구획 조건에는 안정성, 인구기준, 가시적 경계, 형상 조건이 고려되는데, 최적의 기준을 설정하고 여기에서 벗어나는 오차만큼 페널티를 부여하여 페널티가 가장 적은 블록의 조합을 선택한다. 안정성은 CSD, CD 등의 상위 행정단위의 경계를 따르는지의 여부로 결정하며, 인구기준은 최적인구 500명을 기준으로, 형상은 정방형을 최적의 상태로 설정하게 된다. DA의 도입 단계에서는 구획 조건으로 사회·경제적 동질성도 같이 고려하였으나, DA의 빌딩블록인 블록이 2001년 센서스에서에 같이 도입되었기 때문에, 블록별 사회경제적 속성 데이터가 존재하지 않아 실제 DA의 구획에는 이 기준을 고려하지 않았다(Puderer, 2001).

하나의 DA는 최대 99개의 블록의 조합으로 이루어지며, DA의 경계는 통계자료의 원활한 활용을 위해 상위 단위인 CSD와 센서스 트랙의 경계를 따른다. 300명 이상이 거주하는 아파트나 공동주택의 경우에는 DA 내부에 다

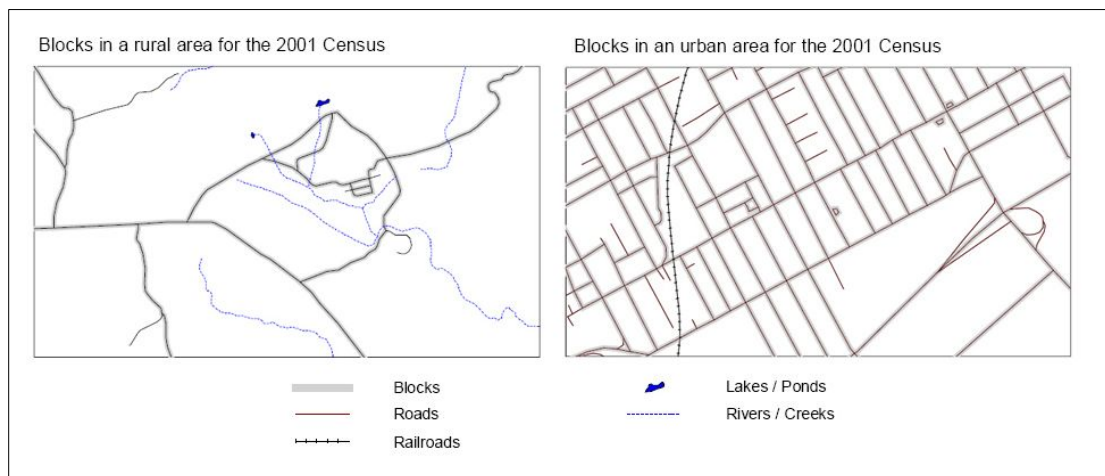
9) 도시 중심부(urban core)의 인구가 50,000 명 이상인 도시 지역, SAC 체계에서 도시 중심부의 인구가 100,000명 이상인 지역은 CMA(Census Metropolitan Areas)로, 중심부 인구가 10,000명 이상인 지역은 CA(Census Agglomerations)로 설정되고, CA 중 중심부 인구가 50,000명 이상인 지역에는 센서스 트랙이 설정되는데, CMA 지역과 CA 중 센서스 트랙이 설정된 지역이 larger urban center에 해당한다(Statistics Canada, 2004).

른 DA가 형성되어 도넛 모양의 DA가 만들어지기도 한다.

DA의 번호는 8자리의 수로 표기하는데 주 번호(2자리) + CD번호(2자리) + DA번호(4자리)로 이루어진다. DA의 번호는 CD 내에서 고유한 네 자리의 수를 부여하는데, S자 형태로 인접한 DA에 연속된 번호를 부여하게 된다.

라. Block

블록은 모든 면이 도로 혹은 표준 지리 지역의 경계에 접하고 있는 기본 단위 공간으로 DA와 함께 2001년 센서스에서 처음 도입되었다. 캐나다의 센서스는 2001년부터 DA를 기본단위로 대부분의 통계 자료가 공표되지만, 인구와 주택 수는 블록을 최소 공표단위로 발표된다. 하지만 개인 정보의 보호를 위해서 인구 15명 이하 블록의 경우에는 반올림해서 자료를 공표하게 된다.



(그림 3-6) 캐나다의 블록 설정

출처 : Puderer(2001), 강영옥외(2007)에서 재인용

5. 호주의 통계구역 설정 현황

가. 호주의 센서스와 통계구역 구분체계

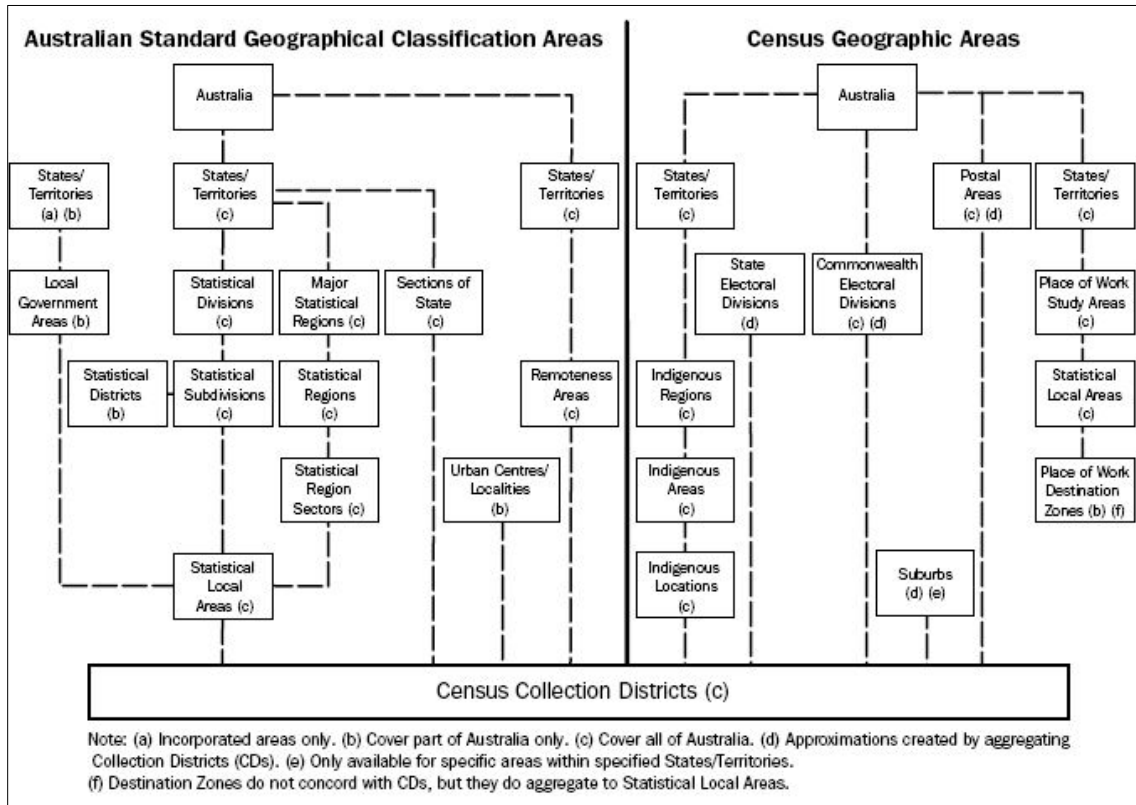
호주의 센서스는 5년마다 이루어지며, 가장 최근의 센서스는 2006년에 이

루어졌다. 호주의 통계구역은 Australian Standard Geographical Classification(ASGC)이라는 체계로 이루어져 있으며, ASGC는 센서스 뿐 아니라 다른 통계 및 지역구분에도 활용되는 체계이다. ASGC는 총 7가지로 이루어져 있으며, 이 중 가장 중심이 되는 것은 센서스의 기본 체계가 되는 Main이다. 지역 구분 체계 중 Main, Statistical Region, Section of State는 전 국토에 설정되어 있으며, Local Government Area, Statistical District, Urban Centre/Locality 의 경우 일부지역에만 설정되어 있다.

Main은 State/Territory - Statistical Divisions(SD) - Statistical Subdivisions(SSD) - Statistical Local Areas(SLA) - Census Collection Districts(CD)의 체계를 갖는다. 통계 자료의 조사 및 공표는 Main의 체계를 따라 이루어지는데, 센서스 자료의 경우 CD를 조사 및 공표 단위로 사용하며, CD는 상위단위인 SLA, SSD, SD 등으로 손쉽게 합쳐진다. 다른 국가의 통계구역과 달리 CD는 센서스가 이루어지는 해에만 설정되기 때문에, 센서스가 이루어지지 않는 해에는 SLA가 가장 최소의 단위가 되며, 센서스 이외의 각종 통계는 SLA를 최소 조사단위 및 공표단위로 사용한다.

나. Statistical Local Area

SLA는 인구주택 센서스 이외의 통계에서 사용되는 통계 자료의 수집과 공표의 기본 단위로, 2006년 현재 1,426개가 설정되어 있다. SLA의 경계는 법률상 정해진 지방자치지역이라고 할 수 있는 Local Government Areas(LGA)가 설정되어 있는 지역에서는 LGA의 경계를 기본으로 하여 설정되는데, 하나의 SSD 내에 위치한 평균적인 LGA는 개별적으로 하나의 SLA로 설정된다. LGA가 설정되어 있지 않은 지역에서는 한 개의 SSD 내에서 충분한 면적이나 경제력을 갖춘 지역이나 섬 등의 경우는 하나의 SLA로 분리하고, 나머지 지역의 경우, 육상 지역(on-shore area)과 해상 지역(off-shore area) 등으로 구분한다.



(그림 3-7) 호주 통계구역(Statistical Area) 체계

출처 : Australian Bureau of Statistics(2007), Trewin(2006)

SLA의 명칭은 하나의 LGA로 이루어진 SLA의 경우 LGA의 이름을 따르며, LGA의 일부분으로 이루어진 SLA의 경우 LGA의 이름에 하이픈으로 추가하거나 지역명을 따르게 된다. 하나의 LGA가 여러 개의 SSD에 걸쳐 있어 분할된 경우에는 LGA의 이름에 Pt A, Pt B 등을 추가한다.

각각의 SLA는 States/Territory 내에서 고유한 4자리 수의 번호를 부여받으며, 하나의 LGA가 SLA로 설정된 경우는 맨 마지막 자리 수를 0번으로, SLA가 한 LGA의 일부일 경우는 1~8번을 부여하며, 비지방자치지역과 해상 지역 등에는 9번을 부여하게 된다. 한 S/T 내에서 일반 SLA는 0001에서 9990번이 부여되며, 특수용도(undefined area) 지역은 맨 뒤 두 자리 번호를 91~99번으로 부여한다.

다. Census Collection Districts

CD는 인구주택 센서스의 조사, 공표의 기본 공간 단위로, 한 명의 조사원이 10일 동안 배포하고 수집할 수 있는 공간적 범위로 설정하는 것이 원칙이다. 하지만 센서스 간 비교를 위해 반드시 이 기준이 지켜지지 않는다는 다른 통계조사의 기본 단위인 SLA와는 달리 CD의 구획은 센서스가 있는 해에만 설정되기 때문에 5년에 한 번씩만 이루어지며, 2006년 센서스의 경우 전국에 걸쳐 총 38,704개의 CD가 구획되었다. CD의 번호는 7자리 수로 이루어지는데, 가장 첫 자리는 주/테리토리의 번호이며, 뒤의 6자리 수는 주/테리토리별로 고유한 번호를 부여한다.

CD는 센서스 자료 수집을 위한 최소 단위로서 공표 과정에서의 개인 정보 보호를 위해 각 CD당 최소 100명 이상의 주민이 포함되도록 구획(원주민 거주 지역은 80명)되며, 2006년 센서스의 경우 도시지역 CD의 평균 인구는 512명이었다. CD의 경계는 상위 공간 단위로 통합이 가능하도록 하기 위해 SLA의 경계를 반드시 따르며, 이전 센서스에서 설정되었던 CD의 경계, Suburb/Locality boundaries, Commonwealth Electoral Divisions(CED)와 같은 통계청 이외의 기관에서 설정한 경계도 가능한 한 유지하는 것을 원칙으로 한다. CD의 경계는 도로, 하천의 중앙과 같은 가시적인 지형지물을 중심으로 설정되는데, 농촌 지역에서는 효율적인 지역 구분을 위해 주요 도로는 가급적 경계로 사용하지 않는다.

라. 메쉬 블록(Mesh Block)

메쉬 블록은 새롭게 도입 준비 중인 통계 자료의 최소 공표단위이다. CD가 센서스 데이터만의 조사 및 공표구역으로 사용되는 것과는 달리 메쉬 블록은 센서스 및 다른 통계 자료의 공표구역으로 모두 이용되는 것을 목적으로 하고 있으며, 조사단위와 분리된 공표단위로서의 기능만을 갖는다. 메쉬 블록은 현재 다양한 목적과 용도를 위해 구분되어 사용되는 여러 계층의 공간 단위와 모두 연동이 가능한 최소 크기의 공간 단위이며, 인구규모는

20~50명 수준이다. 메쉬 블록은 인구 규모면에서 CD에 비해 4~5배 작은 공간 단위이며, CD의 경계를 따르지 않는 독립적인 구역이다. 2006년 센서스에서 시범적으로 처음 도입되었으며 전국에 걸쳐 350,000개가 설정되었는데, 지속적인 검토 및 사례 적용을 통해 2008년 7월 최종 확정하는 것을 목표로 하고 있다. 최종 확정된 메쉬 블록은 2011년 센서스부터 CD를 대체하는 공표단위로 사용될 예정이다.

메쉬 블록의 경계는 GIS tool을 통해 PSMA 호주 국가 DB에서의 해안선, 도로, 철도, 하천 등의 속성, CADLITE, 행정구역 경계, G-NAF 정보, 각 state/territory의 토지이용계획 자료를 이용하여 일차적으로 구획한다. 이후 SLA 경계를 유지하기 위해 SLA 별로 도시/농촌 지역 데이터와 토지이용계획에 따라 인구기준을 조정하고¹⁰⁾, 2006년 센서스에서 메쉬 블록별 인구의 실제 조사를 통해 보정한다.

메쉬 블록의 번호는 S/T 내에서 고유한 11자리의 수로 부여되며, 유지 관리는 1년 단위로 이루어진다. 메쉬 블록의 유지 관리는 경계를 변화시키는 것이 아니라, 인구 증가, 토지 이용 변화 등에 따라 기존의 메쉬 블록을 분할하는 방식으로 이루어진다.

6. 일본의 통계구역 설정 현황

가. 일본의 센서스와 통계구역 구분체계

(1) 국세조사(國勢調査)

일본의 센서스는 국세조사라는 이름으로 5년 단위로 이루어지는데, 10년마다 full-scale census, 사이의 5년마다 short census가 시행된다. full-scale census와 short census는 조사방법이나 범위는 동일하지만 세부적인 조사항

10) 메쉬 블록의 구획은 도시/비도시 지역과 토지이용 형태에 따른 주거지역, 상업지역, 농업지역 등으로 구분하고 인구기준을 조금씩 다르게 설정한다. 예를 들어 도시의 주거지역의 경우, 일반적인 메쉬 블록의 인구기준보다 높은 30-60명 규모로 설정된다.

목에서 차이가 있다. 2차 대전 이전의 센서스에서는 short 형식은 성, 연령, 결혼 여부 등과 같은 기본적인 인구 및 가구 속성만 조사를 하고, full-scale 형식에서 기본 속성 이외에 경제 상태까지 조사하는 체계로 이루어졌으나, 현재는 short census에서도 기본적인 경제 상태 등 세부 항목이 추가된 상태이다. 가장 최근에 실시된 full-scale census인 2000년 센서스의 경우 1995년 센서스 항목에서 ‘현 거주지에서의 거주기간’, ‘5년 전의 거주지’, ‘재학, 졸업 등의 교육 상태’, ‘근로 시간’, ‘통근·통학’, ‘가구 수입 형태’ 등 6가지 항목이 추가로 조사되었다.

센서스의 조사 단위로는 조사구(enumeration district)가 사용되며, 센서스 자료의 공표단위는 기존에 설정되어 있는 행정구역 체계를 그대로 사용하는 데, 전수조사 항목의 경우 광역자치단체인 도(都, To)·도(道, Do)·부(府, Fu)·현(縣, Ken)과 기초자치단체인 시(市, shi)·구(區, ku)·정(町, machi)·촌(村, mura), 그리고 하위의 cho, aza 단위까지 발표되며, 표본조사 항목은 시·구·정·촌 단위까지만 발표된다.

(2) Grid Square Statistics

Grid Square Statistics는 사회경제적으로 증대된 소지역 통계 제공에 대한 요구에 따라 기존의 행정구역 단위로 발표되는 통계 자료 이외에 전국을 경·위도에 따라 0.25-1km² 단위로 구획된 mesh 형태의 통계체계이다. Grid Square Statistics는 통계 자료의 공표 뿐 아니라 도시 계획, 교통 계획, 환경 관리, 상권 분석 등 공공 및 민간영역에서 다양하게 활용된다. Grid Square는 고정적이기 때문에 행정구역의 변화 등에 구애받지 않으며, 시계열적 분석이 용이한 장점이 있다. 센서스를 비롯하여 사업체 통계 등이 Grid Square로 공표되는데, 이는 기존 공표단위로 사용되던 행정구역 단위를 대체하기 위한 것이 아니며, 행정구역 단위의 데이터 공표가 별도로 진행된다.

그리드는 크게 기본 그리드와 분할 그리드, 통합 그리드로 나뉜다. 그리드의 구획은 세 번에 걸쳐 이루어지는데, 우선 위도 40°와 경도 1°단위로 구획하여 1차 그리드를 만들고, 1차 그리드를 다시 8×8로 분할하여 2차 그리드

를 만들며, 이것을 다시 10×10으로 구획하여 기본 그리드를 만든다. 분할 그리드는 기본 그리드를 2×2, 4×4, 8×8로 분할한 그리드를 가리키며, 통합 그리드는 2차 그리드를 5×5, 2×2 또는 분할하지 않은 것을 말한다.

그리드의 번호는 기본 그리드의 경우 8자리의 수를 부여하는데, 앞의 4자리는 1차 그리드를 가리키며, 다음 2자리는 2차 그리드의 번호를 의미한다. 1차 그리드 번호 4자리 중 앞의 2자리는 그리드 최남단의 위도에 1.5를 곱한 값을, 뒤의 2자리는 최서단의 경도값을 나타낸다. 2차 그리드와 기본 그리드 번호의 첫 자리는 상위 그리드 내에서 남쪽에서부터의 순서를, 뒷자리는 서쪽에서부터의 순서를 나타낸다.

나. 조사구와 기본단위구

센서스의 조사 단위인 조사구(Enumeration District)는 조사원이 직접 가구를 방문하여 조사할 수 있는 범위를 기준으로 설정된 공간 단위로, 2000년 센서스의 경우 전국에 걸쳐 940,000개가 설정되었으며, 830,000명의 조사원이 동원되었다. 조사구 당 평균인구는 50명 정도로 설정되며, 조사구의 구획은 인구기준을 충족하도록 기본단위구를 여러 개 결합하는 방식으로 이루어진다.

기본단위구(Basic unit block)는 1990년에 설정된 영구적 기초단위로, 조사구 및 인구집중지역의 설정에 기본단위로 사용된다. 기본단위구의 구획은 법률에 따라 가구(街區) 또는 가구에 준하는 경계를 기준으로 이루어진다. 기본단위구는 보통 3,000~5,000㎡의 면적으로 이루어지며, 주거 표시를 실시하는 구역에서는 하나의 가구가 독립적인 기본단위구로 설정된다. 가구 방식에 의한 주거표시를 하지 않은 구역의 경우는 도로, 철도, 하천, 기타 항구적인 시설 등의 지형지물과 학군, 투표구 등의 지역 구분 등을 따라서 설정한다. 대규모 개발이 이루어지고 있는 지역은 변화 이전까지는 개발지역 전체를 하나의 기본단위구로 설정하며, 이외에 일정 규모 이상의 산악지대, 경지, 저수지 등과 공장, 학교, 철도 및 항만 시설, 병원, 교도소, 군부대 등은 특별구역으로 설정된다.

제3절. 우리나라의 통계구역 및 공표구역 지정 현황

1. 통계권역체계

인구주택센서스 결과의 집계·공표는 행정구역 체계(시도→시군구→읍면동)를 기본으로 이루어진다. 인구주택센서스 중에서도 전수조사 항목의 경우가 기본체계를 따라 공표가 이루어지지만, 10% 표본조사 항목의 경우, 최소단위가 읍면동이 아니라 시군구로 설정된다(2000년 인구주택센서스 기준).

통계자료의 공표는 행정구역 체계를 기반으로 이루어지지만, 통계자료의 수집과 집계에는 행정구역과는 구별되는 별도의 통계지역이 활용되고 있는데, 이러한 통계지역으로 조사구와 기초단위구가 있다. 조사구는 통계자료의 수집을 목적으로 조사원의 업무량을 분배하기 위해 설정된 지역으로, 보통 하나의 조사구는 60가구 정도의 규모를 가지고 있다. 조사구는 인구주택총조사가 실시되는 매 5년마다 설정되며, 설정 당시의 예산, 조사원 수 등에 따라 달라진다.

기초단위구는 기존 통계자료의 수집단위로 이용되었던 조사구가 5년 단위로 설정되면서, 매회 설정을 위한 방대한 인력 및 예산 투입과 같은 문제점이 발생하였던 것을 보완하고, 급증하고 있는 지역통계수요에 부응하며, GIS와 연계한 고급통계작성의 토대를 마련하기 위한 목적으로 2000년 인구주택총조사에 처음 도입되었다. 2003년 10월 현재, 전국에 걸쳐 총 321,174개가 설정되어 있다(통계청, 2003b).

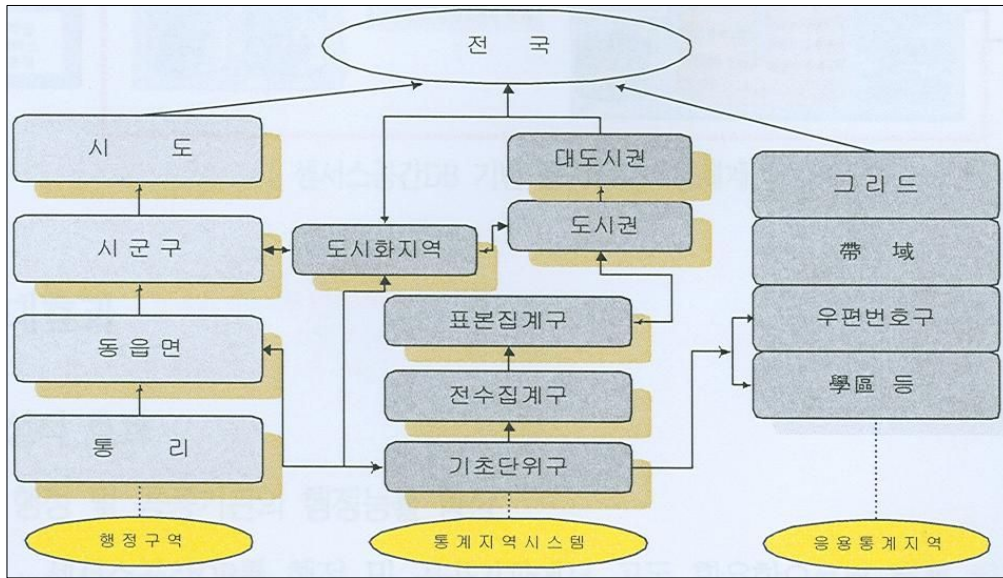
현재의 행정구역 체계에 따른 통계는 권역 설정이 간단하고, 행정정보와 비교가 용이하다는 장점이 있다. 반면에, 지역 내의 다양한 특성을 반영하기 힘들다는 문제를 가지고 있다. 또한 현재 사용하고 있는 통계권역체계의 최소단위인 읍면동 단위가 지역에 따라서는 인구가 2만 명 이상으로 지역 내 동질성을 유지하기에 규모가 너무 크다. 또한, 면적도 1~250km²로 차이가 있다. 이런 이유에서 각종 계획과 사업 시행에서 면밀한 분석이 불가능하다. 더욱이 표본조사의 경우는 이보다 더 넓은 시군구를 공표단위로 하고 있기 때문에 자료의 유용성 측면에서는 더욱 문제가 되고 있다. 다원화되어가는

현대사회의 특징을 정확히 파악하기에 현재의 통계권역체계는 무리가 있으며, 현재의 행정구역 발표단위보다 작은 소지역의 통계 공표단위 도입이 필요하여 새로이 설정되고 있는 것이 집계구이다. 집계구는 공표의 최소단위이며, 집계구의 최소규모는 기초단위구이고, 최대 규모는 행정동이다.



(그림 3-8) 우리나라의 통계권역 체계

이러한 통계권역체계 외에도 현재 통계지리정보과를 중심으로 통계지리정보 구축 및 서비스 관련 사업들이 추진되고 있으며 (그림 3-9), 이들 권역 가운데 주요 권역에 대한 설정방법 및 내역은 다음과 같다.



(그림 3-9) 통계지역체계
출처 : 통계청 통계지리정보과 내부자료

2. 각 통계권역의 개념 및 설정방법

가. 기초단위구

기초단위구는 도로, 하천, 철도, 산 능선 등과 같은 준 항구적인 명확한 지형지물을 이용하여 지도상에 구획한 최소 단위구역으로 2000년 인구주택 총조사에서 처음으로 도입되었다. 기초단위구의 구획은 먼저 읍면동별로 가구수에 관계없이 간선도로, 주요 하천, 산능선 등과 같은 명확한 지형지물을 이용하여 크게 대구역을 구획하고 대구역 내에서 다시 소구역을 구획하는 방식으로 이루어진다. 인구수에 관계없이 정하기 때문에 기초단위구별 인구 및 가구수는 일정하지 않다.¹¹⁾

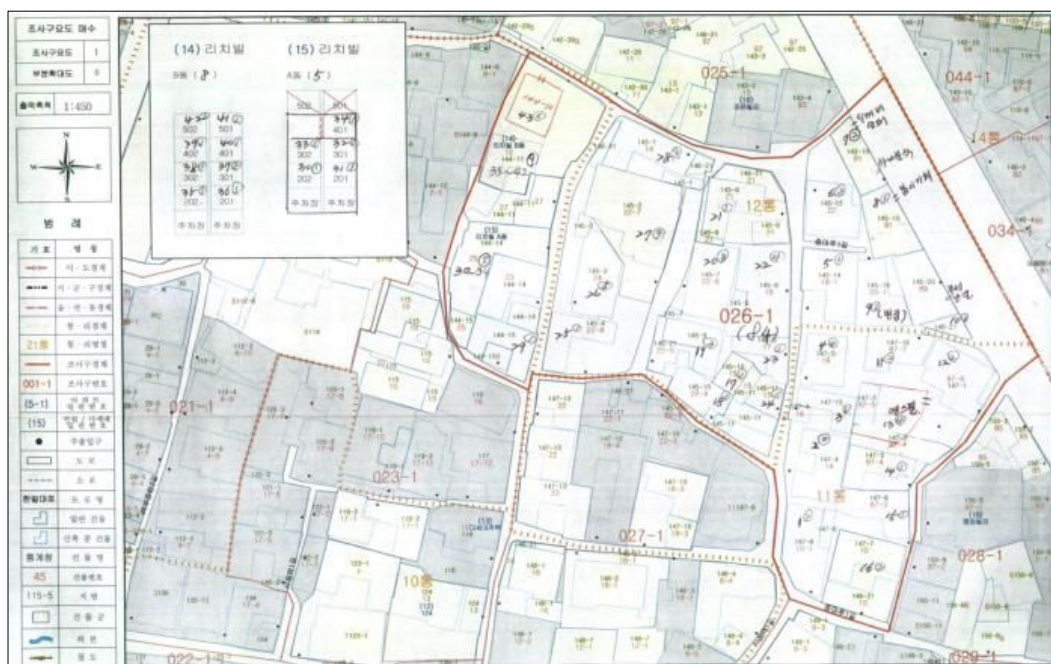
나. 조사구

조사구는 통계자료의 수집을 목적으로 조사원의 업무량을 분배하기 위해

11) 기초단위구에 관한 내용은 <부록 7> 참조

설정된 지역으로, 보통 하나의 조사구는 60가구 정도의 규모를 가지고 있다. 2000년 인구주택총조사에서는 조사구를 군부와 시부로 나누어 단계적으로 설정하였으며, 먼저 읍면동에서 조사구를 가설정 한 다음, 통계청에서 이를 심사하여 조사구를 확정하였다. 군 지역은 다른 지역에 비해 조금 일찍 1999년 9~10월에 조사구가 설정되었으며, 일반 시 지역은 2000년 2~4월에, 특별 및 광역시 지역은 2000년 4~6월에 조사구가 설정되었다.

2000년 인구주택총조사와 달리 2005년에는 읍면동에서 조사구를 설정할 수 없기 때문에 조사구 가설정 및 설정 작업을 모두 통계청에서 수행하였다. 그렇지만 가설정된 조사구와 변동 조사구를 확인하여 조사구를 수정·보완하기 위해서는 지방자치단체의 협조를 받는 것은 불가피하다. 이에 조사구 설정에서는 시군구 통계담당 직원의 업무 협조를 받았다(임명선, 2003). 이런 방식으로 2004년 하반기에 가구 수를 파악하였고, 2005년 상반기에 조사구를 설정하였다.



(그림 3-10) 조사구 예시 도면
출처 : 통계청 내부자료

다. 집계구

우리나라 통계청에서는 현재 통계자료의 최소 공표단위로 행정동 단위(군 지역은 읍·면 단위)를 사용하고 있다 (통계청, 2002). 그러나 행정동이나 읍·면 단위는 적게는 몇 천 명에서 많게는 몇 만 명에 가까운 인구규모로, 하나의 동질적인 지역개념으로 보기에는 여러 가지 어려움이 있다. 사회가 복잡해지고 발전할수록 좀 더 정교하고 세밀한 통계의 작성이 도시 및 지역계획의 수립과 정책의 확립에 필수적인데, 행정동 단위로 광범위하게 공표되는 현재의 통계 체제는 늘어나는 소지역 통계의 수요에 충분히 부응하지 못하고 있다(서울특별시, 2001; 강영옥·이주일, 2005; 통계청, 2003b).

이러한 현실에 입각하여 통계청에서는 2000년 인구 및 주택센서스부터 조사의 기본이면서, 동시에 통계 공표의 기초단위가 되는 기초단위구에 대한 고려를 시작하였다. 하지만 기초단위구는 개인정보 보호나 통계적 유의성 확보 등 통계집계공표구역(이하 집계구)으로 고려해야 할 요소들에 대한 배려가 미흡한 상태에서 지형지물만을 기준으로 경계가 설정되었기 때문에, 그 자체를 집계구로 사용하기에는 여러 가지 한계를 가지고 있다 (통계청, 2003a). 따라서 준영구적인 통계의 기초단위로 설정된 기초단위구를 바탕으로 소지역 통계자료의 공표에 적합한 새로운 개념의 집계구 도입의 필요성이 꾸준히 제기되었으며, 이에 따라 통계청에서는 기존의 행정동·읍·면 단위보다는 작고, 기초단위구보다는 크며, 기초단위구를 기본으로 하는 집계구를 설정하게 되었다.

집계구에 관련된 연구는 2007년 4월 서울시정개발연구원에서 서울시 강남구, 종로구, 용인시, 전주시, 강원도 평창군, 전남 태안군을 대상으로 집계구를 설정하는 작업을 시범적으로 하였으며, 현재 통계청에서 전국에 대해 집계구를 설정하는 작업을 발주하여 2008년 초에 전국적으로 집계구가 설정될 예정이다.

집계구는 기초단위구를 기초로 하여 집계구의 평균인구규모를 500인, 그리고 집계구내의 주택유형과 평균지가라는 측면에서의 사회경제적 동질성을 최대화하고, 집계구의 형상은 면적둘레지수를 사용하여 정방향에 가까운 형

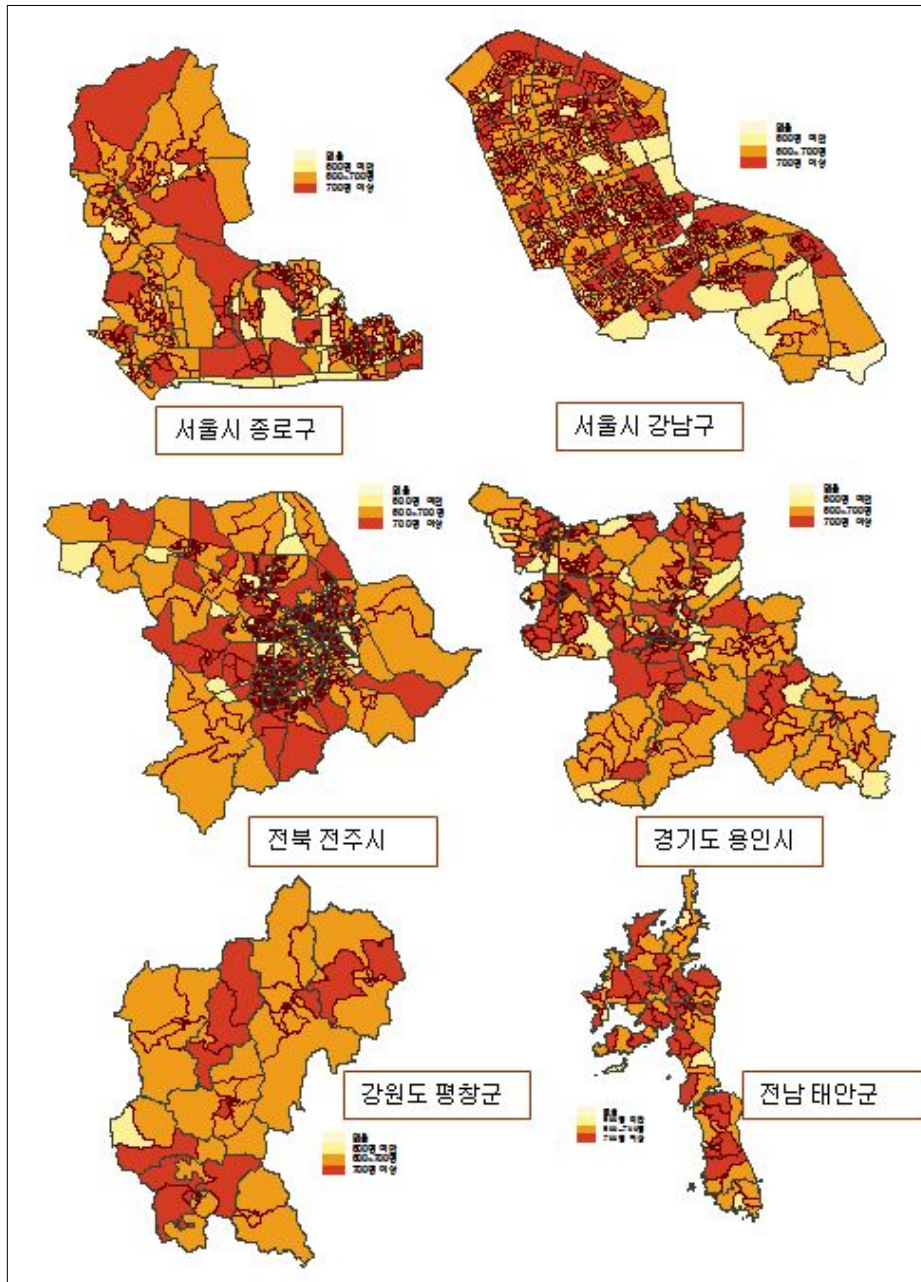
태를 유지하도록 하여 설정하였다. 이러한 다양한 요소를 고려하고 최적의 기초단위구들의 조합을 찾기 위해 AZP(Automatic Zoning Procedure) 프로그램을 개발하였으며, 각 요인별로 가중치는 인구: 주택유형: 지가: 형태간 0.4: 0.3: 0.1: 0.2로 설정하여 사례지역에 대한 집계구를 설정하였다. (그림 3-11)은 6개 지역에 대한 집계구의 예다.

라. 도시화지역

1995년부터 도시지역과 농촌지역이 통합된 도농복합 형태의 시가 나타나면서, 종전에 행정구역을 토대로 도시지역으로 간주되어온 5만 이상의 시지역(또는 2만 이상의 읍지역 포함)의 성격이 모호해졌다. 외국에서 일반적인 행정구역과는 별도로 인구밀도 및 토지이용기준을 활용하여 도시지역을 정의해서 확정하고 있으며, 도시화지역 설정도 이러한 필요성에 의해 시작되었다.

2004년과 2006년 수행된 연구(통계청, 2004c, 2006b)에서 도시화지역은 통계청의 기초단위구를 도시화지역 확정의 기본단위로 활용하였으며, 첫째, 인구밀도 3,000명/Km² 이상이거나 지목으로 정의된 도시적 토지이용 면적 비율이 1/2을 초과하는 기초단위구를 도시화지역 확정대상에 포함시켰다. 둘째, 두 기준 중 최소한 하나를 만족하면서, 공간상으로 접해 있는 복수의 기초단위구와 그러한 기초단위구들로 둘러싸인 내부의 기초단위구들로 구성된 군집을 도시화지역 후보지로 추출하였다. 그 다음에 이들 지역 중 인구 3,000명 이상인 지역을 도시화지역으로 최종 확정하는 기준을 설정하였다.

이러한 도시화지역 설정기준을 서울특별시와 경기도의 시·구를 대상으로 실제 적용하였으며, 도시화지역 설정을 위한 시스템을 개발하여 향후 도시화지역의 설정과 유지, 관리에 활용하도록 하였다. 현재 통계청에서는 전국에 대한 도시화지역 설정 용역사업이 발주된 상태이며, 2008년 초까지 전국에 걸친 도시화지역 확정이 완료될 예정이다.



(그림 3-11) 사례지역별 집계구 설정결과
출처 : 강영욱외(2007)

제4절. 원자료 활용 증진 방안

1. 원자료 수요 증가와 자료 서비스 문제

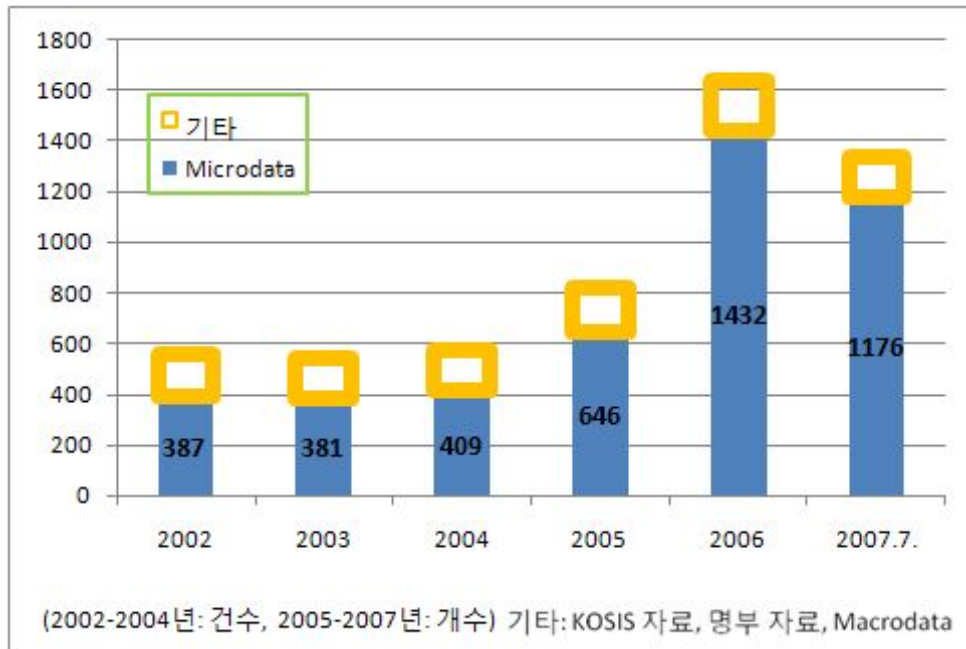
원자료에 대한 관심은 주로 세밀한 연구에서 발생한다. 일반적인 정책 관련 작업이나 민간부문의 마케팅 등에서는 적절한 수준의 집계자료를 이용하는 것이 편리하다. 원자료를 다루기 위해서는 원자료를 가공하는 데 많은 시간과 노력 및 비용이 따르기 때문에 빠른 시간 안에 마쳐야 하는 사업에서는 실제로 원자료를 이용할 수 없다. 또한 원자료를 이용하기 위해서는 원자료를 가공할 수 있는 다양한 능력과 기술이 필요하고, 이를 분석할 수 있는 통계처리 능력이 필요하다. 그러므로 원자료 이용에 대한 요구는 이러한 능력을 가진 전문 인력이 증가하면서 함께 증가하게 된다. 이런 측면에서 보면, 우리나라의 고학력 증가에 따른 전문인력의 증가는 원자료에 대한 수요 증가로 나타나는 것이 지극히 당연한 일이다.

지난 10여 년에 걸쳐 정부와 관련 있는 연구기관이나 대학이 증가하였다. 이와 함께 지방자치단체가 설립하여 운영하는 연구소도 양적으로 많아졌을 뿐만 아니라 그들의 연구 활동도 증가하였다. 이러한 현상은 중앙정부나 지방정부의 운영에서 과학적인 근거를 바탕으로 정책을 수립하고 집행하는 과학적 행정이 점차 확대되어 가기 때문에 나타난다.

한편, 민간부문에서도 기업들이 커지고, 때로는 세계 시장을 선도하면서 연구역량도 크게 증가하였다. 이와 함께 시장의 흐름을 판단하고 미래전략을 세우기 위해 기업 산하 연구소들도 빠르게 성장했다. 특히, 세계화의 물결로 국제 교류가 커지고, 외국 자본이 들어오면서 한국 시장에 관한 단기적 연구도 양적으로나 질적으로 모두 크게 성장하였다.

이러한 연구의 성장은 집계 수준의 자료만으로 파악하기 어려운 세밀한 연구로 이어진다. 예를 들어, 정부에서 생활보호대상자를 선정하는 기준을 마련하는 경우 각 가구의 가구구성원별 소득을 알아야 하고, 경제활동상태를 알아야 하며, 이를 가구별로 집계해야 한다. 민간기업의 예를 들자면, 기업이 어느 지역에 대리점을 내고자 할 경우, 어느 위치가 적절한지를 판단하기 위

해서는 읍면동 수준보다 작은 지역 수준의 집계자료를 가져야 한다. 이러한 작업들을 위해서는 원자료를 연구자가 자신의 필요에 맞게 가공하여야 한다. 이런 이유에서 최근 원자료 이용이 크게 증가하고 있다((그림 3-13)).



(그림 3-13) Microdata 이용자 증가 현황

출처 : 통계청 전산개발과 통계자료제공 현황(통계청 내부자료)

이와 더불어, 지금의 시대적 조류 또한 원자료 제공을 압박하고 있다. 세계화의 흐름에서 국가경쟁력은 한 국가가 반드시 지녀야 할 요소이며 자원이 되었다. 국가 간 경쟁에서 이전처럼 국가가 국가의 경쟁력을 이끌어 가는 역할은 한계에 이르렀다. 이제 국가는 민간 기업들이 더욱 강한 경쟁력을 가질 수 있도록 촉진하는 역할을 한다. 이런 측면에서 최근 Web 2.0과 비슷하게 Gov 2.0이란 용어가 한창 주목을 받기 시작했다. Web 2.0은 어떤 특정 전문가가 아니라 일반 사람들이 만들어내는 정보(contents)가 중요한 정보의 자원이 되는 것을 의미한다. 이와 비슷하게, Gov 2.0은 정부가 일반 사람들이 좀 더 정확하고 질 좋은 정보를 생산할 수 있도록 기초를 마련하는 것을 의미한다. 이런 차원에서, 통계청은 가공하기 이전의 통계자료를 일반 사람

들에게 보급하여 그들이 직접 가공한 통계자료를 바탕으로 자신에게 필요한 정보를 생산할 수 있도록 함으로써 우리 국민의 정보역량을 높일 수 있다. 여기에는 물론 통계역량도 포함되어 있다.¹²⁾

이렇듯 역동적인 시대적 상황에서, 통계청의 원자료 제공에는 몇 가지 해결해야 할 문제가 있다. 예전에 비해 전반적으로 통계청의 서비스는 아주 많이 향상되었다. 그렇지만 아직도 작고 세심한 부분에서 서비스를 받지 못하는 경우가 종종 있다. 특히 특정 목표를 지향하는 인과분석, 연구 목적에 초점을 맞춘 자료가공 등 세밀한 자료처리에서는 이러한 불편이 많이 나타나고 있다. 이러한 현상은 연구 인력의 증가와 함께 더욱 많아지고 있으며, 이에 따라 불만의 수준도 높아지고 있다. 이러한 이유는 크게 통계 수요자 및 통계 자료에 따라 제도적으로 차별화된 서비스가 필요한데 그렇지 못하기 때문이며, 다른 하나는 통계자료에 관해 객관적으로 접근이 가능하면서도 상세히 자료를 설명하는 문건이 매우 부족하기 때문이기도 하다.

여기서는 이런 문제들에 대한 해결 방안을 찾기 위해 먼저 해외의 원자료 서비스에 관한 전반적 경향을 살펴보고, 다음으로는 우리나라의 원자료 제공과 관련된 문제를 구체적으로 찾아보고, 이의 해결을 위한 방향을 제시하고자 한다.

2. 원자료 서비스의 전반적 경향

우리와 마찬가지로 선진국에서도 원자료에 대한 수요 증가는 그리 오래된 현상이 아니다. 이에 최근 미국과 유럽의 각국들을 중심으로 원자료 제공에 대해 많은 발전이 있었다. 특히 ICT 기술의 발달에 힘입어 사용자들이 먼 거리에서 원자료에 접속하여 분석까지 가능하도록 해주는 서비스로 확대되고 있다.

12) 이의 가장 유명한 예는 아마도 현재 우리나라의 TV 시청률을 발표하는 Nielson의 경우일 것이다. 1980년대 Nielson은 미국의 센서스 자료를 이용하여 Nielson Television Rating을 만들었으며, Nielson Food and Drug Indexes도 만들었다(Hawkers, 1986). 이렇게 Nielson은 센서스 자료를 바탕으로 index 중심의 모형을 개발하여 세계 각국에서 컨설팅서비스로 큰 사업을 벌이고 있다.

그러나 원자료 서비스의 확대는 한편으로는 많은 사람들이 편리하게 자료를 접할 수 있도록 하는 장점이 있지만, 원자료는 개인의 정보를 담고 있어 이에 대한 보호가 더욱 중요해졌다. 이런 이유에서 많은 국가에서는 원자료에서 개인정보가 드러나지 않도록 자료를 가리거나(masked), 그러한 정보를 삭제(removed)하거나, 여러 경우를 합친(collapsed) 다음 원자료를 제공하고 있다. 그렇지만 특정 연구에서는 개인정보를 가린 자료를 이용하고는 연구 성과를 낼 수 없기 때문에 이러한 자료에 대한 서비스 방안도 마련해 놓고 있다.

다시 말해, 원자료의 이용자가 증가에 따라 이용하는 원자료의 종류에 따라 차별적 서비스를 제공하고 있다. 먼저 누구나 이용할 수 있는 원자료를 Public Use Microdata(이하 PUM)라 부른다.¹³⁾ 이러한 자료는 가능하면 모든 사람이 손쉽게 이용할 수 있도록 하고 있다. 최근에는 온라인으로 자료를 다운받을 수 있게 하고 있으며, 대체로 무료로 배포하고 있다. 이러한 자료도 예전에는 주로 CD와 같은 매체를 통해 제공했으며, 그 이전에는 학교나 연구기관 등과 협조하여 마그네틱 테이프 등으로 배포하였다.

한편 개인정보가 담긴 원자료(Confidential data)에 대해서는 예전부터 상당히 까다로운 이용 절차를 거쳐 특정 이용자들만 이용할 수 있도록 해 왔다. 이용에서도 통계청이나 통계청이 지정한 장소에서만 분석을 할 수 있도록 on-site 서비스를 하였으며, 자료를 지정된 장소 밖으로 유출할 수 없고 분석 결과만 가지고 갈 수 있도록 정하여 자료를 제공하였다. 이러한 규정은 그다지 크게 달라지지 않았다. 다만 최근에는 좀 더 발달된 ICT 기술을 이용하여 원격접속(remote access 또는 remote analysis)으로 연구 장소에서 자료에 접속하여 통계적 분석을 수행하고 분석결과는 받을 수 있게 하는 국가가 늘어나고 있다. 이러한 노력들은 모두가 이용자들이 이용의 편의성을 높여 좀 더 좋은 통계정보를 생산하기 위한 것이다.

13) 과거에는 개인수준의 자료를 raw data(원시자료)라고 불렀으나, 최근에는 microdata라고 부르는 것이 보편적이다.

3. Public Use Microdata

선진국들은 다양한 방식으로 Public Use Microdata를 서비스해 왔다. PUM의 가장 대표적인 것이 센서스 자료다. 통상 전수조사인 센서스 자료에서 일부를 추출하여 제공하는데, 미국의 경우는 1%와 5% 표본을 제공하고 있으며, 우리나라는 2% 표본을 제공하고 있다. 이렇게 표본을 추출하여 제공하는 이유는 센서스 자료가 너무 방대하기 때문이기도 하지만, 이렇게 함으로써 개인정보를 보호하기 수월해지기 때문이다.

PUM을 제공하는 방식은 국가마다 또 기관의 성격에 따라 제각각으로 매우 다양했다. 예를 들어, 미국의 경우를 보면 센서스 원자료를 배포하는 방식이 다양하다. 센서스국(Census Bureau)의 홈페이지에 접속하여 다운받을 수 있으며, CD나 DVD로도 받을 수 있다. 또한 ICPSR (Inter-University Consortium for Political and Social Research)이나 IPUMS (Integrated Public Use of Microdata Series)를 통해 이용할 수도 있다.¹⁴⁾

센서스 원자료는 단지 한 국가에서 뿐만 아니라 국가 간 비교에서도 중요한 역할을 한다. 이런 취지에서, 1999년 10월에 출범한 IPUMS-international은 세계 각국의 센서스 자료에 관한 목록을 작성(inventory)하고, 센서스 자료를 서로 비교가 가능하도록 조정(harmonize)하며, 이렇게 만들어진 자료를 배포(disseminate)하고((그림 3-14) 참조), 없어질 위험이 있는 원자료와 관련 문건을 보존(preserve)하는 역할을 한다(Hall et al, 2000). 현재 IPUMS에는 전 세계에서 68개국이 참여하고 있으며, 26개 국가의 80개 센서스 원자료를 제공하고 있다. 아래 (그림 3-15)는 IPUMS의 분포를 표시한 것으로, 진한 녹색은 센서스 자료를 함께 쓰고 있는 국가, 그보다 옅은 녹색은 자료 비교 가능하도록 협조하는 국가, 아주 옅은 녹색은 협상이 진행 중인 국가이며, 노란색과 보라색은 아직 참여하지 않는 국가다. 참고로 한국 통계청은 아직 참여하지 않고 있으나 참여를 고려 및 검토하고 있다.¹⁵⁾

14) ICPSR은 1962년에 설립한 대학의 연합체로 미시간(Michigan)대학에 본부를 두고 있다. IPUMS는 미네소타 인구연구소가 1999년 10월에 설립한 기관으로 IPUMS-international의 중심 역할을 하고 있다.

15) Choe(2000)에 하와이대학 동서문제연구소가 보유한 한국의 센서스 목록이 실려 있다.

4. Confidential data

가. 북미

미국은 Center for Economic Studies(CES)라는 센서스국 산하 기관을 통해 정부, 학계, 민간의 연구를 장려하고 지원한다. 이를 위해 CES는 미 전역에 9개의 Research Data Center(RDC)를 운영하고 있다. RDC가 있는 곳은 대체로 큰 도시의 대학 및 연구소 또는 센서스국 관련 기관이다. RDC가 있는 9개 지역은 Boston의 National Bureau of Economic Research, Berkeley 대학, Los Angeles 대학, Washington D.C.의 CES, Chicago 대학, Michigan 대학, New York의 Baruch 행정학교, Cornell 대학, Duke 대학이다.

Confidential data에 대한 접근을 RDC 안에서만 가능하다. 다시 말해 개인정보가 담긴 원자료를 이용하는 경우 RDC 밖으로 결과만 가지고 나올 수 있을 뿐이며, 자료를 가지고 나올 수는 없다. 더욱 까다로운 절차는 일단 RDC에서 자료를 이용하는 권한을 얻는 데 있다. 자료를 이용하기 위해서는 연구계획서를 제출하고 심사과정을 거쳐 CES의 승인을 받아야 한다. 승인절차는 최소한 3개월이 걸리는 것으로 되어있지만, 통상 6개월에서 12개월이 걸린다. 승인을 받는 데는 인력, 연구 가능성, 학문적 기여도, 이용의 정당한 사유가 중요한데, 특히 센서스국의 프로그램에 유익함을 보여주어야 한다. 이를 위해 진술서(Benefit Statements)를 제출하며, CES는 이와 관련된 13개의 기준을 제시하고 있다(CES, 2003).

연구 과정에서도 정보보호에 대한 안전성의 규제는 매우 심각하다. 연구의 중간과정에서 만들어지는 결과물(intermediate output)에 대해서 이를 RDC 밖으로 가져온다든지 세미나 등에서 발표를 할 때도 제한을 두고 있다. 특히, 논문을 쓰거나 연구조교를 하는 대학원생의 만들어내는 중간결과물에 대해서는 지도교수가 RDC에 와서 검토하도록 하고 있다.

CES의 이용자에 대한 서비스는 기능중심이다. 먼저, 연구계획서 제출이나 진술서 제출 등에서도 RDC 이용의 가능성, 현실성 등을 위해 RDC의 담당자와 상의할 것을 권하고 있다. 일단 심사를 통과하면 RDC는 연구자들이

confidential data를 이용하는 데 연구공간을 제공하고 다양한 형태의 서비스를 제공한다. 또한 결과물 산출 후 발표 등에서도 규정위반 등의 위험을 피하기 위해 RDC의 담당자와 상의 할 것을 권하고 있다.

<표 3-3> 미국의 RDC 운영 현황

설치도시 (설치기관)	Cambridge (NBER)	LA (UCLA)	Chicago (FRBC)	Ann Arbor (미시건대)	Ithaca NY (코넬대)	Durham (듀크대)
설립년도	1994	1999	2002	2002	2006	2000
관계 연구기관	NBER	UCLA	FRBC, Chicago Univ. Illinois Northwestern Univ. Argonne National Laboratory	미시건대학 ICPSR	Baruch Coll. NY City U. FRB Fordham U. NBER NY U. Pace U. Princeton U. Russell Sage Foundation, Rutgers U.	Duke U. N. CarolinaU NC State U
센서스국 직원 수	1명 (경제학박사)	2명 (1명 경제학박사)		1명 (경제학박사)		1명 (경제학박사)
직원 외 스텝수	0명	0명		2명 (대학교원)		1명 (듀크대학교수)
부스의 수	8개	단말기 8개		8개		8개
이용시간	24시간	24시간		24시간		24시간
이용요금	\$15,000/년	\$3,125/월, \$15,000/년		\$15,000/년, 미시건대학교원, 대학원생무료	비참가기관은 \$20,000/년	\$15,000/년, 관계대학교원, 대학원생무료
연간 수락 proporsal 수	12개			15개		5-6개
이용자 중 대학원생비율	약 50%	약 33%		약 50%		약 42%

연구 후의 절차도 까다롭다. 최종결과물에 대해서는 과정보다 더욱 심각한 규제가 있다. 최종결과물도 RDC의 사전 심사를 받아야 외부로 가지고 갈 수 있으며, 연구결과를 발표하는 데도 허가를 받아야 한다. 또한 연구결과에서 센서스국에 유익한 내용을 문서로 만들어 CES에 제출하여야 하며, 연구내용은 working paper로 제출하여야 한다. 만약 어떤 방법으로든 이런 절차를 피해 정보보호를 위반하는 경우가 적발되면 징역 5년 또는 25만 달러 이하의 벌금형을 받을 수 있다.

한편, 캐나다의 confidential data 제공은 미국과 유사하다. 다만 캐나다의 경우 미국과 비슷한 방식으로 전환된 것은 그다지 오랜 일이 아니다. 예전에는 캐나다 microdata 자료를 이용하기가 어려웠기 때문에 연구자들이 상대

적으로 비용이 싼 미국의 microdata를 이용하여 연구를 하였다. 이에 대응하여 캐나다는 최근 Data Liberation Initiative를 추진, DLI에 참여하는 교육기관에는 연간비용을 부담하고 무제한으로 public use microdata를 이용하게 하였다. 또한 confidential data에 대해서도 미국과 비슷한 이용방식을 채택하여 20개소의 caRDC를 운영하고 있다. 특징은 미국에 비해 교육이나 학위를 목적으로 대학원 학생들의 이용을 적극적으로 추진하고 있다. 이용자는 상당히 많은 편이어서 2006년 기준 연간 776개의 프로젝트, 이용자 853명(연구자 516명, 학생 337명)이다.

나. 유럽

유럽에서도 개인정보가 포함된 자료를 이용하고자 하는 요구가 증가하고 있다. 이에 Eurostat은 2002년에 개인정보가 담긴 confidential data의 이용에 관한 규정을 두어 European Community 자료를 이용할 수 있도록 하고 있다.¹⁶⁾ 여기에는 EC 소속 국가들이 보유한 confidential data의 상호이용에 관한 규정도 함께 포함되어 있다. 대학, 과학적 연구 수행 기관, 그에 준한 기관, Committee on Statistical Confidentiality의 승인을 받은 경우, EC나 소속 국가의 행정부서로부터 의뢰를 받은 과학적 연구를 수행하는 기관 등이 confidential data를 이용할 수 있다.

미국과 마찬가지로 연구계획서가 중요한데 과학적 기준에 맞는 상세한 연구계획서가 있어야 하며, 연구제안서에 구체적인 이용 자료의 내역, 분석 방법, 이용 기강이 명시되어 있어야 한다. 계약에 이용조건, 연구자의 의무, 비밀유지의 구체적인 방법 및 이의 위배 시 처벌에 관한 사항이 명시되어 있어야 하고, 이에 대한 연구자와 소속기관의 서명이 필요하다. 연구에서도 EC가 정한 범위 안에서 정부의 감독을 받아야 하며, 연구결과를 유출하거나 발표 또는 출판 경우도 confidentiality에 관한 점검을 거쳐야 한다.

북미의 경우와 마찬가지로 전반적으로 confidential data 이용에 대해 심각

16) Commission Regulation (EC) No 831/2002 of 17 May 2002 implementing Council Regulation 322/07 on access to confidential data for scientific purposes.

한 규제가 있으며, 비밀유지에 대한 의무 및 책임이 명시되어있고, 이용승인 및 결과 발표 절차가 까다롭다. 다만 명문화 할 경우 강조점이 조금씩 다를 수는 있다. 또한, 유럽은 몇몇 국가들이 on-site 서비스에 remote access나 remote analysis 같은 방식으로 전환하여 이용편의를 도모하고 있다.

네덜란드의 경우는 미국의 RDC와 비슷한 Research Center Microdata (CeReM)를 운영하여 on-site 서비스를 제공함과 동시에, 원격에서 접속하여 (remote access) 분석을 하거나 분석프로그램을 작성하여 보내면 실행 (remote execution)하여 결과를 받을 수 있는 서비스도 함께 시행하고 있다. 이용자들을 위해 SAS, SPSS, STATA, OX 등 다양한 통계패키지를 제공한다. 원격접속에서는 본인임을 확인하기 위해 지문인식과 PKI 방식의 인증방식을 사용하고 있다. 몇 가지 흥미로운 점은 원자료 사용은 무료지만, 원자료 서비스에는 비용이 따른다는 것이다. 이 비용에는 원자료에 대한 메타자료 생성, 이용에 대한 상담, 연구과정 및 최종결과 점검에 대한 비용 등이 포함되어있다. 또한 원자료를 이용한 연구과제가 네덜란드 통계청의 임무와 상충되어서는 않된다.

핀란드는 1990년대 증가한 연구로 2001년 Research Laboratory라는 on-site를 공식적으로 운영하기 시작하였다. 1990년대 중반부터 시작한 confidential data를 공식화한 것이어서 이미 연구자들 사이에는 이러한 제도가 잘 알려져 있었다(Statistics Finland, 2006). 핀란드의 원자료 이용에 관한 전반적인 절차는 다른 국가와 비슷하지만 핀란드 통계만의 독특한 상황 때문에 confidential data 이용에서 특별한 조건들이 강조된다. 핀란드 통계청은 모든 통계와 행정자료를 연결하여 생산하기 때문에 거의 모든 부문에서 개인수준의 원자료를 생성해낼 수 있다. 따라서 핀란드 통계청은 통계청의 자료와 다른 행정자료를 연결하여 연구자가 필요로 하는 원자료를 생성해주기도 한다. 이때 전국이나 특정지역 전체의 개인별 자료는 이용할 수 없다고 명시하고 있다. 또한 특정 집단 전부에 대한 개인별 자료에 대해서도 아주 특별한 경우가 아니면 이용할 수 없다. 더 나아가, 원자료 제공시 연구자가 이미 보유하고 있는 원자료도 점검한다. 왜냐하면 기존의 자료와 새로운 자료를 연계하여 개인정보를 생성할 위험이 있기 때문이다.

덴마크는 세계 최초로 행정자료로 통계를 생산하기 시작한 국가다. 이러한 역사적 배경에서, 덴마크는 1979년 최초로 원자료 보호법을 제정하였으며 이 법은 2000년 개인정보의 처리에 관한 법(Act on Processing of Personal Data)으로 개정하였다. 또한 1980년대 중반에 이미 원자료에 수요가 나타나, 통계청에서는 1986년 on-site 서비스를 제공하기 시작하였다. 2001년 통계청과 연구 및 연구환경부(Ministry of Research and the Research Environment)는 원자료에 대한 연구자의 이용을 촉진하기 위해 통계청에 연구서비스 조직을 신설하고, 연구실에서 통계청 원자료에 원격접속이 가능하도록 하였다(Borchsenius, 2006). 이때 원격접속은 특정 장소로 한정되어 있기는 하지만, 요건을 갖춘 특별한 경우에는 집에서 접속할 수 있도록 하고 있다. 그렇다고 모든 자료에 대해 원격접속이 승인된 것은 아니다. 기업자료와 같이 비밀보호가 어려운 자료는 원격접속을 할 수 없다. 또한 이용하는 자료는 표본자료여야 하며, 만약 전수일 경우는 변수에 대한 제한을 두고 있다.

제5절. 소결

우리나라에서 인구주택총조사의 이용을 확대하는 것은 그다지 손쉬운 일이 아니다. 행정에서 인구주택총조사가 이용될 수 있는 데는 행정단위를 결정하거나 선거구를 획정하는 업무인데, 여기에는 모두 주민등록자료가 이용되기 때문이다. 그렇지만 이러한 제약에도 불구하고 주민등록으로 파악할 수 없는 사회조사 성격의 총조사에 대한 지방자치단체들의 수요가 커지고 있다. 이와 더불어 민간부문과 학술 연구에서도 총조사에 자료에 대한 수요가 늘고 있다. 이 가운데 가장 대표적인 것이 소지역통계이며, 총조사의 표본조사에 대한 소지역통계에 대해서는 더욱 그러하다. 또 다른 수요는 원자료에 대한 것이다. 이런 측면에서 이 장에서는 통계구역과 원자료 서비스에 관해 다뤘다.

외국의 통계구역 사례들을 바탕으로 현재 진행되고 있는 우리나라의 통계구역설정을 살펴보았다. 그 가운데 특히 소지역통계 제공에 중심 역할을 하

는 집계구를 중점적으로 살펴보았다. 집계구와 더불어 기초단위구는 도시화 지역과 도시권 확정에 사용될 수 있다. 특히 현재 도농복합 지역이 발생하여 행정구역을 정의하는 데 어려움이 있는 지역들이 있다. 이런 지역에서 통계에 바탕을 두고 행정구역을 나눈다면 좀 더 현실적이며 효율적인 행정의 공간을 마련할 수 있을 것이다.

집계구는 통계를 가능한 가장 작은 지역 수준까지 생산함으로써 지방자치단체가 많은 정책에 통계를 직접 활용할 수 있도록 만들어 준다. 이러한 통계는 단지 행정뿐만 아니라 민간기업의 마케팅, 입지선정, 표본조사의 표집틀 등으로 다양한 활용성을 잠재하고 있다. 또한 국가 중심이었던 학술적 연구가 지역별로 비교연구로 확대될 수 있는 가능성을 제공한다.

한편 이러한 작업들은 모두 GIS의 기술적 지원이 또 다른 한 축을 지탱하고 있다. 다시 말해, 전국을 작은 지역 단위로 나누어 그것을 통계의 목적으로 재결합하는 작업은 GIS의 지원이 없다면 구현하기 매우 어려운 것이다. 이런 측면에서 여기서는 인구주택총조사의 활용성을 높이기 위해서는 GIS 기술을 기반으로 전국을 작은 지역으로 나누고, 이를 통계적 동질성과 공간적 근접성으로 묶어 통계와 지역이 함께 짊어질 수 있게 만들 것을 권한다. 이렇게 함으로써, 마치 개인의 정보가 개인의 행동을 예측하는 데 사용되듯이 소지역의 통계는 그 지역에 대한 정책, 그 지역 주민들의 행위를 예측하는 등 여러 가지 용도로 사용될 수 있을 것이다.

또 다른 활용에 대한 제안은 원자료 이용을 활성화하는 것이다. 일단 우리나라 통계청의 통계 서비스는 매우 빠르게 성장해왔음을 긍정적으로 받아들이면서, 그 가운데 원자료에 관한 부분이 상대적으로 뒤떨어졌으며, 더욱이 그에 대한 수요가 다른 무엇보다 빠르게 성장하고 있다는 사실에 주목하였다. 현재 기술적인 차원에서 원자료 서비스는 MDSS와 CD와 같은 매체를 통해 어느 정도 이루어지고 있다. 또한 on-site 서비스도 작은 규모로나마 이루어지고 있다. 그렇지만 아직은 이에 대해 뚜렷한 정책의 방향이 정립되지 않았다.

여기서는 먼저 Public Use Microdata와 Confidential data를 구분하여, 차별적 서비스를 권한다. 거의 모든 선진국들이 이와 같은 방식으로 원자료를

서비스하고 있다. 또한, PUM에 대해서는 무료로 제공하고, Confidential data에 대해서는 절차와 범규를 까다롭게 적용하고 비용을 책정하고 있다. 이는 일반적 수요와 전문적 수요를 차별화하여 각각에 적절한 방식으로 대응하는 것이다. 더 나아가, 원자료 이용의 증가는 특화된 정보생산의 증가를 의미한다. 이것은 지금과 같은 정보화 시대의 국가 간 경쟁에서 국가가 국민들의 정보역량을 높이는 데 필요한 기초 자원을 제공하는 것이다. 이런 의미에서 이제 통계청은 통계생산이라는 작은 범주에서 벗어나 국가의 가장 기초가 되는 정보를 생산하고, 이의 유통 및 활용을 장려하여, 국가 경쟁의 초석으로 역할을 바꾸어 나아가야 할 것이다.

제4장. 행정자료 기반의 센서스

제1절. 서론

현재 우리나라의 많은 국가통계는 조사를 바탕으로 작성된다. 조사를 바탕으로 하지 않는 국가통계는 행정적 절차를 이행하는 과정에서 만들어지는 자료를 바탕으로 작성된다. 예컨대, 주민등록통계이며, 범죄통계, 국세통계 등도 이러한 범주에 속한다. 그렇지만 인구주택총조사나 산업총조사와 같은 국가통계의 가장 기본이 되는 통계나 사업체기초통계조사, 임금구조기본통계조사, 국민건강실태조사, 한국복지패널 등과 같이 다른 통계의 기본이 되는 통계는 거의 전적으로 조사에 의해 작성되고 있다.

우리나라와 비슷하게 외국도 대체로 조사를 바탕으로 국가통계를 작성하고 있다. 그러나 최근 들어 단독가구의 증가로 조사대상을 접촉하기 어렵게 되고 응답에 대한 거부감이 증가하여 조사환경이 나빠지고 있으며, 더불어 조사비용의 증가로 전통적인 조사방법으로 국가통계를 작성하기가 점점 어려워지고 있다. 특히 인구센서스와 같은 전수조사에서 이러한 어려움이 나타나고, 무응답이 많으면 참으로 대처하기가 어렵다. 실제로 이런 경우가 유럽의 일부 국가에서 나타났다. 대표적인 예가 네덜란드와 독일로, 네덜란드는 1970년대, 독일은 1990년대 이후 인구센서스를 중단한 상태이다(Bierau, 2000). 이러한 변화에 대응하여 선진국들은 다양한 국가통계생산 방식들을 모색하고 있으며, 특히 전수조사에 의존해왔던 인구센서스에 대해 적극적으로 새로운 조사방식을 모색하고 있다.

선진국과 마찬가지로 우리나라에서도 조사환경이 빠른 속도로 악화되고 있다. 더욱이 우리의 경우 읍면동사무소 기능축소로 말미암아 과거 실제 조사에 도움을 주었던 행정지원이 없어 앞으로 조사의 어려움은 더욱 커질 것으로 보인다. 이런 변화에 대응하여 우리나라에서도 2000년대 초부터 행정자료를 국가통계생산에 활용할 수 있는 방안을 모색하기 시작하였다(김형석, 1999; 홍두승외, 2002; 이삼석외, 2002; 김형석외, 2003; 이건의외 2004).

우리나라는 행정자료 중심의 국가통계 작성과 관련하여 몇 가지 중요한 기반을 갖추고 있다. 첫째, 1962년부터 시행한 주민등록제도가 있다. 주민등록제도는 모든 국민 개개인을 파악할 수 있는 주민등록번호가 있을뿐더러, 40년 이상 제도화되어 오류가 거의 없는 편이며, 더 나아가 모든 행정 및 공식 기관에서 동일한 주민등록번호를 사용하고 있다. 둘째, 높은 정보화 수준과 2000년부터 시작된 전자정부사업으로 행정자료를 연계·통합할 수 있는 기반이 잘 갖추어져 있다.

이런 근거를 바탕으로 최근 인구주택총조사를 주민등록으로 대체하자는 의견이 빠르게 확산되고 있다. 이 장에서는 행정자료 기반 센서스의 이론적 가능성과 모형을 살펴본다. 이를 위해 외국에서 시행되고 있는 행정자료 기반의 센서스도 함께 살펴본다.

제2절. 외국의 센서스

1. 다양한 센서스 방식

전통적인 전수조사의 센서스 방식에 대한 문제는 유럽에서 나타나기 시작하였다. 가장 큰 어려움을 경험한 국가는 독일과 네덜란드다. 독일은 1983년과 1987년 센서스의 경험 때문에 2000년에 센서스를 실시하지 않았으며, 2011년 센서스라운드¹⁷⁾에 참가를 기뻐하고 있다. 네덜란드에서도 1971년 센서스에서 응답거부율이 0.026%의 경미했으나, 1981년 센서스에서는 예상 응답거부율이 26%에 이르러 전통적인 방법의 센서스를 하지 못하였다.

이러한 어려움은 1970년대 중반 이후 선진사회에서 서서히 개인정보와 사생활에 대한 보호의 관점이 강화되면서 나타나기 시작했다. 국가와 개인 사이에서 개인의 중요성이 부각되고 센서스를 국민의 사생활에 대한 국가의 부당한 간섭으로 생각하는 사람들이 생겼다. 이에 따라 센서스에 대한 국민들의 비난과 거부, 보이콧 운동, 헌법소원 등이 있었다. 국민들의 거부의 직접적인 원인은 조사가 사생활 침해라는 인식과 이를 증폭시킨 통계자료의

17) 구동독에서는 1981년에 마지막으로 센서스가 실시되었다.

행정활용이었다. 특히 센서스에서 파악된 정보를 행정에서 활용할 수 있는가는 첨예한 쟁점의 대상이었다.

이런 과정에서 1979년 덴마크에서 통계와 개인정보에 관한 법이 처음 만들어진 이후, 1980년대를 거치면서 여러 나라에서 통계에서 개인정보를 보호하는 법들이 만들어졌다. 이에 통계자료의 행정활용은 국민들에게 정부의 사생활 간섭이라고 받아들여지고, 법적으로 “정보의 사용에 대한 자기결정의 원칙”이 생겨나, 조사된 개인별 통계자료가 행정의 목적으로 활용할 수 없게 되었다. 이때부터 센서스에서 자료를 이용하기 위해서는 이름을 가리고, 주소의 공개수준을 제한하는 등의 간단한 방법부터 좀 더 복잡한 자료 마스킹 기법들이 개발되기 시작하였다.

센서스에 대한 거부가 확대되면서 1990년대에는 2000년 센서스라운드 이후 센서스에 대한 논의가 활발해져, 다양한 방식의 센서스가 논의되고 있다. 그러한 논의들 가운데 일부는 이미 실행에 옮겨져 점검을 하는 중이거나, 이미 어느 정도 실행에 옮겨진 것도 있다. 이들을 센서스의 실행 방식과 계획에 따라 다음과 같은 네 범주로 분류할 수 있다(<표 4-1>).¹⁸⁾

<표 4-1> 센서스 방식에 따른 국가 분류

전통적 방식	과도기	행정자료 활용	새로운 방식
그리스	벨기에	덴마크(완료)	네덜란드(혼합)
스페인	룩셈부르크	핀란드(완료)	독일(혼합)
아일랜드	스위스	아이슬란드(진행)	프랑스(순환)
영국	오스트리아	노르웨이(진행)	미국(전수·순환)
이탈리아		스웨덴(진행)	싱가폴(혼합)
포르투갈			

출처 : Laihonen(1999), ONS(2003), 西村善博(2006)에서 재구성

18) 홍두승·이건·이명진·서우석·장원호(2002) 연구에서 활용한 자료들과 분석결과들을 이용하여 이루어졌다. 독일 통계청에서는 이와 조금 다른 방식으로 분류하고 있다. 이에 관해서는 <부록 2> 을 참조하기 바란다.

가. 전통적 방식

전통적 방식을 사용하는 국가에서는 현지조사를 통한 기존의 센서스 방식이 계속 유지되고 있다. 행정자료는 현지조사를 조직하고 자료를 수합하는 과정에 주로 활용된다. 행정자료의 활용은 센서스를 위한 보조적인 수단이라 할 수 있다. 그리스, 이탈리아, 포르투갈, 스페인과 같은 남유럽 국가들과 영국, 아일랜드가 이 범주에 속한다. 이들 국가들에서는 기존 센서스 방식을 행정자료 활용으로 대체하기 위한 계획이 아직 없으며, 기존의 센서스 방식을 유지한다는 점에서 공통된다. 그러나 스페인, 영국과 같은 일부 국가에서는 기존의 센서스 방식을 유지하면서도 센서스와 관련하여 행정자료를 더욱 폭넓게 활용하고 있다. 한 가지 흥미로운 사실은 스페인이 2001년 센서스에서 세계 최초로 모든 응답자들에게 인터넷 응답을 하도록 허락했다는 주장이다 그러나 인터넷 응답에 대해서는 잘 알려지지 않았으며, 이용률이 그다지 높지 않았던 것으로 보인다(ONS, 2003: 6).

나. 과도기

과도기에 있는 국가들은 전통적인 방식의 센서스를 장기적으로 더 이상 실시하지 않고 행정자료를 활용하는 방식을 도입하려 하고 있다. 오스트리아는 2001년 센서스를 끝으로 전통적 방식의 센서스를 더 이상 하지 않기로 했고, 벨기에는 행정자료 활용 방식과 전통적 방식을 모두 활용할 예정이며, 룩셈부르크는 1999년 가을에 행정자료 활용의 센서스 방식에 대한 시험을 하였으며, 스위스는 설문지를 활용하는 조사를 하고 있으나 앞으로 행정자료를 활용할 계획이다.

다. 행정자료

센서스에서의 행정자료 활용은 스칸디나비아 국가 등 주로 북구의 국가들에서 시작되었다. 스칸디나비아 국가들에서는 1960년대부터 지역 등록 체제

를 바탕으로 인구등록의 중앙관리 방식이 도입됨으로써 기본적인 인구통계들이 단일한 자료로 통합관리 되었다. 1970년대부터 센서스 중 인구통계 부문에 행정자료가 활용되기 시작하여 행정자료 활용의 범위가 점차 확대되었다. 1981년에 덴마크, 1990년에 핀란드에서 인구주택 부문과 고용부문 통계를 포괄하여 전적으로 행정자료에 의존하는 센서스가 완성이 되었다. 반면 스웨덴과 노르웨이에서는 행정자료의 활용이 아직 부문적으로만 이루어지고 있다. 스웨덴에서는 1985년에 인구부문에 더하여 고용부문 행정자료를 통합하였으나, 주택부문 자료는 아직 전통적 방식으로 조사하고 있다. 노르웨이는 스웨덴과 유사한데, 1990년에 표본센서스 실험을 하였으며, 행정자료 활용 확대가 지연되고 있는 상황이다. 아이슬랜드는 인구주택부문을 행정자료를 통해 조사하고 있으나, 경제활동에 대한 자료는 통합되지 않았다. 이들 국가, 특히 덴마크와 핀란드의 경우는 아래에서 좀 더 상세히 다룰 것이다.

라. 새로운 방식

프랑스는 시의성과 응답자 부담의 분산을 염두에 두고 순환형 센서스로 전환하여, 1999년 이후부터는 일종의 상시적인 센서스를 시행하고 있다. 미국은 2010년 센서스에서 전수조사부문은 그대로 시행하며, 표본조사 부분은 시행하지 않을 계획이다. 표본조사는 American Community Survey라는 순환조사 방식으로 대체하여 매년 300만 가구를 표본조사하여 10년 동안의 누적치를 사용하게 된다. 프랑스와 미국의 경우는 뒤에 순환형 센서스 부분에서 다시 상세히 다룰 것이다.

네덜란드는 1971년이 마지막 센서스를 시행한 후 1981년과 1991년에는 가용한 행정자료와 표본조사를 이용해 센서스 형식의 자료를 생산했다. 이를 가상센서스(virtual census)라 부른다. 네덜란드는 현재는 순수히 Eurostat에 자료를 제공하는 목적으로 가상센서스(Virtual Census)를 시행하고 있다(ONS, 2003). 독일도 네덜란드처럼 센서스의 어려움을 겪은 후 현재는 매년 1%의 작은 센서스(microcensus)를 통해 센서스 자료를 갱신하고 있다. 문제는 표집틀이 1987년 센서스에 기초해 있어 이를 해결하기 위해 행정자료를

이용하는 모형이 시험 중에 있다(ONS, 2003). 한편 멀리 동남아시아의 싱가포르는 2000년에 행정자료로 센서스를 대체했다. 이와 함께 20%의 표본조사를 실시하여 행정자료로 얻을 수 없는 자료, 교육, 고용 등을 확보하였다. 표본조사는 인터넷으로 하였으며, 후속 조치로 전화, 다음은 방문을 하였다.

2. 행정자료의 보완적 활용

가. 영국의 One Number Census¹⁹⁾

영국은 1801년 이래 매 10년 마다 센서스를 실시해왔다. 조사원이 가구를 방문하여 설문지를 배부하고, 응답자는 기입하여 임시지역사무소에 우편으로 부치게 되어 있다. 우편으로 송부되지 않은 경우, 조사원이 직접 수합을 한다. 이와 같은 방법에는 특별한 것이 없으나 누락을 추정해내는 데 영국은 독특한 방법을 사용하고 있다. 이를 One Number Census(ONC)라 부른다. One Number Census는 2001년 센서스에서 32만 가구(1.48%)를 표본추출하여 Census Coverage Survey를 실시한 다음, 이를 바탕으로 지역별 인구자료를 보정하여 인구수를 공표하였다.

보완된 통계치의 신뢰성을 검증하기 위하여 여러 행정자료를 활용하고 있다. 행정자료는 보건자료(Health Authority in England and Wales, Northern Ireland Central Health Index, Community Health Index in Scotland), 영국 군 자료, 미군 및 미군부대 종사자 자료, 사회복지 수혜 아동 자료, 연금수령자자료, 교육부처에서 제공되는 학생 자료 등을 포괄한다. 행정자료에서 추출된 행정정보는 집계 수준에서만 사용되며 주로 특정 인구집단의 센서스 추정치와 비교된다.

나. 네덜란드의 Virtual Census

네덜란드는 1981년 센서스가 사생활 보호 관련으로 국민의 호응이 감소하

19) ONS(1998)에서 정리

여 연기되면서 그 이후 센서스를 시행할 수 없었다(Vliegen and van de Stadt, 1989). 그렇지만 1981년과 1991년에 센서스와 같은 자료를 생산하였는데 이는 행정자료를 활용한 것이다(ONS, 2003). 이때 부분적으로는 표본조사를 실시하였으나 이 둘 사이의 문제로 행정자료의 집계만 이용하였다. 그 이후 행정자료 활용의 법적 근거가 마련되어 행정자료를 충분히 활용할 수 있게 되었다.

1996년 네덜란드는 개인수준에서 가용한 모든 정보를 통합한 데이터베이스, SSB(Social Statistisch Bestand)를 구축하기 시작했다(ONS, 2003). SSB의 구축단계는 3단계로 나뉘어 있다(Ai and Bakker, 2000; Al and Altena, 2000; Arts, 2000). 1단계는 예비조사(pilot study)로서 성공적으로 수행되었으며, 2단계에서는 1995년 시점의 가용한 자료들을 바탕으로 SSB의 원형이 구축되었고, 3단계에서는 1995년부터 1998년 사이에 생산된 모든 가용한 자료들을 모아 자료통합을 하고 있다.

SSB에 사용되는 자료는 다음의 세 가지 종류로 구분된다. 먼저 행정자료는 기본적 인구등록자료(Population Register), 고용보험자료, 사회 공공부문 장애인혜택자료, 소득정보, 부동산 평가자료, 지리기초정보 등을 포괄한다. 다음으로 기업자료는 표본조사를 통해 고용 및 소득에 관한 연간 자료를 확보한다. 가구자료는 표본조사인 노동력 조사, 가구생활비 조사, 사회경제패널 조사, 삶의 질 조사 등에서 확보한다(Gaasemyr, 1999; Schaafsma-Harteveld, 1999; Slagter, 1999, van de Laan, 2000). 이 가운데 인구등록자료가 SSD의 가장 중요한 근간(backbone)을 이루고 있다.

그간의 경험을 바탕으로 2001년 센서스에서 SSD는 충실히 자신의 역할을 하고 있다. 그렇지만 여전히 행정자료에서 얻을 수 없는 정보를 표본조사자료에서 얻고 있다. 이런 이유에서 자료들 간의 일관성을 확보하는 방법들, 특히 반복가중(repeated weighting)법이 개발되어 적용되고 있다. 앞으로 Eurostat이 요구하는 좀 더 세부적인 통계를 생산하기 위해서는 소지역추정 기법(small area estimation)의 개발이 필요하다(Statistics Netherland, 2006; Nordhlt and Gircour, 2004).

다. 독일의 Microcensus와 행정자료

독일에서는 지난 1987년 센서스를 마지막으로 전통적 방식의 센서스를 실행하지 못하다가 지금은 행정정보를 활용하는 센서스를 대안으로 삼고 있다(Statistisches Bundesamt, 1992). 1980년대 이전에 독일을 센서스 결과를 행정자료를 갱신하는 데 사용했다. 이에 대해 1983년 법원은 센서스의 이중적 이용을 금하였다(Juristisches Internetprojekt Saarbrücken, 2001). 이러한 법적 변화에서 1987년 센서스는 진행되었는데, 크게 대중의 호응을 얻지 못했다. 그 이후 독일정부는 센서스 시행을 승인하지 않게 되었다.

일반 센서스의 부재로 microcensus는 큰 역할을 해왔다. Microcensus는 인구와 노동시장에 대한 1% 표본조사로 센서스 사이의 간극을 메우고자 1957년부터 시작되었다. 현재는 2004년에 수정된 “인구와노동시장및가구의주택상황을표본조사시행에관한법”에 의해 지속적으로 시행되고 있다. 이법은 한시적으로 2005년부터 2012년까지 적용된다. 1%의 표본을 조사하기 때문에 매년 대략 39만 가구의 83만 명을 조사하게 된다. 법에서 응답의 의무를 부과해 2005년 microcensus에서는 무응답률이 4%로 작은 편이다(Statistisches Bundesamt Deutschland, 2007).

한편, 전통적 방식의 센서스를 대체하기 위해 활용하게 될 행정자료는 지자체(Gemeinde)에서 관할하고 있는 거주지 등록 자료, 노동청에서 관리하는 직업관련 자료, 주택과 건물의 소유주에게 서면조사를 통해 수합될 자료 등의 세 가지 종류로 구분된다. 수합된 자료는 신뢰도 조사를 거쳐, 가구단위로 통계 자료가 수합이 되어 통합된 센서스 자료로 완성될 예정이다.

현재 행정정보를 활용한 센서스의 본격적 도입을 위해 실행한 센서스 시험조사 Zensustest 결과를 분석중이다. 시험조사의 실행을 위해 “센서스 시험조사를 위한 법”이 2001년 7월 공포되었으며 이 법에 따라 조사가 실행되었다. 시험조사는 2001년 12월 5일 실시되었다. 자료가 일정 기간 이전의 상황을 반영하기 때문에, 조사시점의 현황을 파악하기 위한 재수합이 실시되었다. 지자체 자료는 1개월 이후 1차 재수합이, 4개월 이후 2차 재수합이 이루어졌고, 노동청의 자료는 7개월 이후 재수합이 이루어졌다. 시험조사의 자료

분석에 18개월이 소요되며 시험조사 결과의 일차 발표는 2003년이다. 시험조사의 범위는 독일 내 지자체 Gemeinde의 4%(약 14 500 개중 최대 570개)와 건물의 0.2%(약 1640만 개 중 최대 38,000개)이었다. 시험조사에서는 표본으로 추출된 건물과 인구에 대한 직접 조사의 결과가 행정자료를 통해 수합된 자료와 비교 분석하였다(Statistisches Bundesamt, 2002).

이러한 노력을 바탕으로 현재 연방과 주 통계국에서 여러 행정자료를 바탕으로 센서스 통계를 생산하는 여러 모형이 검토되고 있다. 그 가운데 가장 유력한 모형으로 연방모형(Federal model)과 지방자치단체 모형(Länder' model)이 결합하였다(ONS, 2003). 연방모형은 개인 수준의 자료 결합 없이 소지역의 인구를 행정자료로 만드는 방법이고, 지방자치단체 모형은 개인 수준에서 자료를 결합하는 모형이다. 현재 2011년 센서스에서는 절충적 안이 만들어졌다. 센서스 날에 지방정부는 인구등록자료를 통계청으로 보낸다. 이때 모든 등록자료가 아닌 일부를 발췌하여 보낸다. 이것은 중복등록을 점검하기 위한 것으로, 점검은 개인의 주요 특성 성, 생일, 출생지, 출생국, 성, 태어난 때 이름, 이름 등을 이용한다. 단, 중복점검의 결과는 다시 지방정부에 보내주지 않는다.

이와 함께 같은 날 표본조사가 계획되어 있다. 이는 전 국민의 10%를 대상으로 하는 표본조사다. 이 조사는 행정자료로 집계한 자료의 오류를 점검하는 데 사용된다. 또 건물과 주택 소유주에게 보내는 건물과 주택에 대한 우편조사가 2주 전에 발송되어 센서스 날의 상황에 대해 응답하게 된다. 이 자료와 지방정부의 행정자료 및 연방고용청의 자료를 바탕으로 논리적으로 절차(algorithm)를 적용하여 거처, 가구, 고용, 거주관계를 찾아내어 가구를 구성한다. 여기에 표본조사 결과조사가 이용된다. 이렇게 구성된 가구를 바탕으로 인구와 가구를 집계한다는 계획이다(Statistische Ämter, 2007).

라. 스페인

스페인의 행정정보 활용은 지자체 Municipales와 통계청 Instituto Nacional de Estadística 간의 긴밀한 협조관계를 바탕으로 이루어지고 있다.

스페인에서는 지자체와 통계청이 행정자료를 공동관리하고 있다. 법령에 따라 지자체가 행정자료의 변동사항을 매달 통계청에 보고함으로써 과거에 5년에 한번 씩 갱신되었던 행정등록 정보가 최근에는 상시적으로 갱신되게 되었다(Nacional de Estadística, 1999a).

상시적으로 갱신된 행정자료에서 파악된 행정정보는 센서스에서 직접적으로 활용된다. 행정자료에 의해 파악할 수 있는 문항들은 센서스 설문지에서 제외된다. 대신 응답자들은 센서스 설문지 외에 자신의 행정자료가 제시된 서류를 받게 되고 잘못된 정보를 수정할 수 있는 기회를 갖게 된다. 이를 통해 응답자들의 입장에서는 응답 부담이 줄게 되며, 행정적 측면에서는 행정정보의 정확성과 신뢰성을 높일 수 있게 된다. 상시적으로 갱신되는 행정자료를 바탕으로 장기적으로는 통시적 연구 자료를 발전시킬 계획이다(Nacional de Estadística, 1999b).

3. 행정자료 기반의 센서스

여러 방식 가운데 행정자료 기반의 센서스는 많은 나라들이 궁극적으로 추구하는 센서스 통계의 생산방식이다.²⁰⁾ 그런 의미에서 행정자료 기반으로 센서스를 시행하는 국가들에 대해서는 좀 더 상세히 다루고자 한다. 노르웨이와 스웨덴은 아직 완전한 행정자료 기반의 센서스를 시행하지는 않지만 긴 기간을 준비해 왔기 때문에 그 시행과정에서 우리는 시사점을 찾을 수 있다. 덴마크와 핀란드는 이미 오래 전부터 행정자료로 센서스 통계를 생산하고 있어 우리에게 많은 시사점을 줄 수 있다. 주로 살펴볼 대상은 행정자료 관리 체계, 특히 고유번호의 관리 체계로 이를 바탕으로 우리가 가야할 방향을 찾아보고자 한다. 이와 함께 그러한 행정자료 관리 체계를 도입하기 까지 역사적 과정을 간략히 살펴봄으로써 우리가 거쳐야 할 과정과 그에 따

20) 2007년 12월 3일부터 6일까지 이집트 카이로에서 “2010 World Programme on Population and Housing Censuses: Workshop on Civil Registration and Vital Statistics in the UNESCWA Region”으로 UN 통계국(Statistics Division) 주관으로 센서스에서 행정자료의 보완적 활용에 관한 국제회의가 있는 것으로 보아도 행정자료 기반의 센서스가 중요한 쟁점임을 알 수 있다.

르는 추진 과정에서 필요한 것들을 찾아보고자 한다.

가. 노르웨이

노르웨이는 1769년 처음 인구센서스를 시작하여, 1801년에는 이름과 거주자의 상세내역을 함께 조사하기 시작하였다. 1890년 이후부터는 매 10년에 한 번씩 센서스를 시행해왔다. 한편 인구등록(Population Register)은 1964년에 설립되어 모든 국민이 개인번호(Personal Identification Number)를 갖게 되었다. 1970년부터는 성, 나이, 혼인상태, 가족관계, 국적, 거주지 등의 상세 정보는 인구등록에서 가져다 사용하기 시작하였다. 이때 개인에 관한 정보를 갱신하고 행정에 사용할 명부를 작성하였다. 1980년에는 행정정보를 이용하여 일부 조사항목을 대체하였으며, 조사에 우편조사방법을 도입하였다. 그 이후 등록자료에서 가져오는 정보는 점점 많아졌다. 1990년 센서스에서 인구의 주택에 관한 정보 가운데, 고용과 직업, 통근·통학, 가구구성, 주거환경은 조사로 확보해야 했다. 이를 위해 약 28%의 표본조사를 했다.

2001년 센서스에서는 인구, 교육, 노동시장, 소득 관련 자료는 행정자료에서 가져올 수 있을 만큼 행정자료가 좋아졌으나, 이에 대한 검증이 필요했다. 특히 사업체의 위치에 대한 확인을 위해 사업체등록에 정확한 주소를 도입하여 토지, 주소, 건물 등록(Ground Property, Address, and Building, 일명 GAB register)과 연계가 가능해졌고 검증 또한 가능해졌다. 또한 1999년에는 조사를 바탕으로 해외에서 학위를 취득한 고학력자의 등록도 갱신하였다.

그러나 이러한 노력에도 거처에 대한 등록이 없었기 때문에 인구등록에 있는 개인을 거처와 연결시킬 수가 없었다. 따라서 거처와 가구 부분은 조사에 의존할 수밖에 없었다. 따라서 2001년 센서스에서도 주택부분에 대해서는 조사를 해야 했다. 한편 2000년 국회(Storting)예산심의에서 거처별로 고유한 주소를 GAB Register에 등록하기 위한 거처주소프로젝트(Dwelling Address Project)를 승인하였다. 이에 거처주소프로젝트와 주택센서스가 상호보완할 수 있도록 센서스를 한해 늦추어 2001년 11월에 실시하게 되었다. 이제

노르웨이는 거쳐주소프로젝트가 완료되면 완전히 행정자료로 센서스를 대체할 수 있을 것이라 기대하고 있다. 노르웨이는 아직 완전히 행정자료 기반 센서스는 시행하지 못하지만 30년 동안 꾸준히 행정자료 기반의 센서스로 진화해 온 모습을 보여준다.

<표 4-2>는 노르웨이 센서스에서 주요 항목을 행정자료로 대체한 연도를 표시한 것이다.²¹⁾ 주택과 건물 관련 항목에서는 아직도 전적으로 조사에 의존하거나, 행정자료와 조사결과를 함께 사용해야 하는 항목들이 많다. 한 가지 주목할 만한 것은 아파트(flat)에 관한 것이다. 2001년 조사에서 아파트의 개별 거처에 모두 번호를 부여했고 이를 인구등록과 주소등록에 모두 올려놓았기 때문에 아파트 거처는 모두 식별이 가능해졌다.

<표 4-2> 노르웨이의 항목별 행정자료 대체 연도

시작 연도	항목	시작 연도	항목	시작 연도	항목
	인구관련	1990	사업체	●	편의시설
1980	성별	-	사회경제지위	●	거주기간
1970	혼인상태	1960	소득		건물관련
-	모국어		교육관련	●	건축기간
1970	국적	1990	교육정도	●	건축재
-	종교		가구관련	2001	건물의 주용도
	경제관련	●	가구종류/인원	●	주거단위의 수
1990	경제활동상태	●	가족종류/인원	●	난방시스템
1990	종사상 지위		주택관련	2001	건물의 좌표
1990	산업	●	주택 크기	-	복도면적
-	직업	●	방의 수	●	건물의 층수

- : 원래 공표하지 않는 항목

● : 아직 행정자료로 대체하지 못한 항목

출처 : Laihonen(1999)과 Utne(2005)에서 재구성

나. 스웨덴

21) 여기 제시한 항목은 핀란드 센서스 항목을 기준으로 하였다. 왜냐하면 핀란드가 행정자료로 가장 많은 항목의 통계를 생산하기 때문이다.

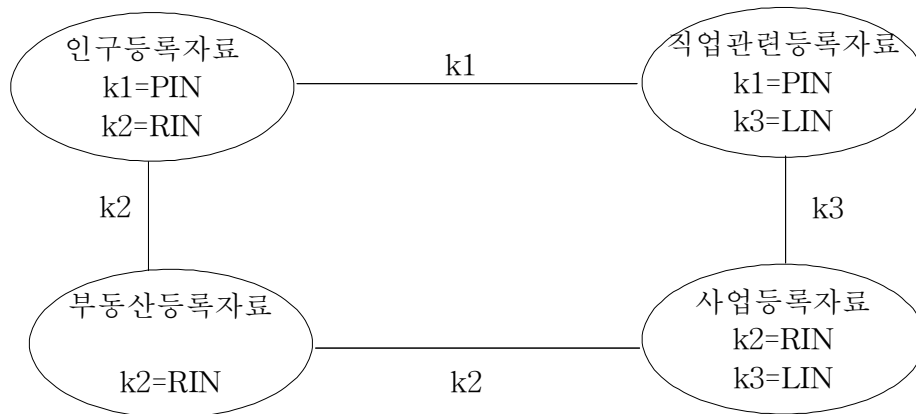
스웨덴이 센서스에서 행정자료를 이용하기 시작한 것은 1950년부터다. 이것은 스웨덴이 오랜 인구등록의 체계를 갖고 있었기 때문에 가능했다. 오래 전 현대적 국가 이전에 교구(parish)에서는 주민을 파악하기 위해 주민을 등록하였다. 거의 450년 전인 1571년이다. 또한 스웨덴 통계청은 1749년에 체계적인 국가통계의 관리를 위해 설립되었으며, 1860년도에는 최초로 인구, 주택 센서스가 실시되었다. 이런 역사적 배경에서, 개인식별번호(personal identification number)가 비교적 이른 1947년에 도입되어 행정자료기반 통계 생산을 위한 기본적인 기반이 형성되었다.

1968년에는 개인식별번호를 포함하여 체계적인 행정자료인 TPR(Total Population Register)이 구축되었다. 1998년도에는 기존의 정보를 보완하여 새로운 형태의 TPR이 구축되었다. 새로운 TPR은 개인식별번호, 이름, 주소, 부동산, 교구, 출생지, 시민권, 가족관계, 수정, 발생시점 등 정보를 포함하고 있다.

기존방식 센서스는 1990년에 마지막으로 실시되었다. 무응답의 증가가 가장 큰 원인이었다. 수도인 스톡홀름 지역에서는 10% 정도, 기타지역에서는 2.8% 정도의 사람들이 응답을 거부했다. 이러한 무응답 비율과 함께 빠르게 증가하는 센서스비용으로 인해 1995년 스웨덴 국회에서는 행정자료를 이용한 센서스를 실시할 것을 결정했다. 이를 바탕으로 스웨덴 통계청에서는 2000년도 센서스를 행정자료로 완전히 대체하려 하였으나 그렇게 되지는 않았다. 2005년에 행정자료로 완전 대체를 선언하였다(Bruhn).²²⁾

스웨덴의 행정자료 기반 통계는 크게 네 개의 상호연계적인 행정자료로 구성되어 있다. 여기에는 인구등록자료, 부동산등록자료, 직업관련 등록자료, 사업관련 등록자료 등이 포함되어 있다. (그림 4-1)은 이러한 구성체계를 도표로 보여주고 있다.

22) 그 결과는 아직 잘 알려지지 않고 있다.



(그림 4-1) 스웨덴 행정자료 기반 센서스 구성도

K1: PIN (Personal Identification Number, 개인인식번호) -인구등록자료
 K2: RIN (Residence Identification Number) 또는 거처번호(Dwelling Identification Number) - 거주 및 건물관련 행정정보
 K3: LIN (Local Identification Number: Workplace Number,

행정자료를 연계하기 위해서는 무엇보다도 공통된 코드를 갖는 것이 중요하다. 그러나 현재 스웨덴에서는 거처를 코드화한 RIN가 누락되어 있다. 거주지 코드 대신 주소를 사용하고 있으며, 모든 주거지에 고유번호 부여작업이 실시되고 있는 중이다. 1983년 이전에 지어진 건축물에 대해서만 번호 부여가 완료되었다. 특히 주거지 고유번호 등록을 과세업무와 관련하여 강제적으로 사용하여 거주번호사용을 촉진하고 있다.

각 행정자료의 유지와 보수는 정보의 종류와 특성에 맞게 정기적으로 실시하고 있다. 가장 기본적인 인구등록 자료의 경우를 살펴보면, 월별로는 개인식별 번호, 거주지, 성별, 연령, 시민권, 출생지, 출생국, 소득 등 정보를 갱신하고 있다. 분기, 년별로는 성명, 주소, 출생교구, 출생국(외국인 경우), 이민여부, 가족관계, 부모 출생국 등 정보를 유지, 관리하고 있다.

그러한 결과 주거에 관한 항목 이외에는 대체로 행정자료로 대체할 수 있게 되었다. 항목별로 행정자료로 대체한 연도를 <표 4-3>에 담았다.

<표 4-3> 스웨덴의 항목별 행정자료 대체 연도

시작 연도	항목	시작 연도	항목	시작 연도	항목
인구관련		1985	사업체	•	편의시설
1950	성별	•	사회경제지위	•	거주기간
1970	혼인상태	1950	소득	건물관련	
-	모국어	교육관련		•	건축기간
1950	국적	•	교육정도	-	건축재
-	종교	가구관련		•	건물의 주용도
경제관련		•	가구종류/인원	•	주거단위의 수
1985	경제활동상태	1985	가족종류/인원	•	난방시스템
-	종사상 지위	주택관련		1985	건물의 좌표
1985	산업	-	주택 크기	-	복도면적
•	직업	•	방의 수	-	건물의 층수

- : 원래 공표하지 않는 항목

• : 아직 행정자료로 대체하지 못한 항목

출처 : Laihonen(1999)에서 재구성

다. 덴마크

(1) 행정자료 기반 센서스 자료 생산 방법

덴마크 통계청은 행정자료를 크게 개인, 주거 및 건물, 사업장에 관한 부분으로 나뉘어 파악하고 있다. 개인에 관한 정보는 개인의 이름, 주소, 주소 변경, 가족관계, 출생, 사망, 결혼, 이혼, 이민, 입양, 시민권 취득, 인구등록 날짜 등으로 중앙인구등록에 포함되어 있다. 중앙인구등록은 내무부 주민등록청(Ministry of Interior, the Central Office of Civil Registration)이 총괄하여 관리하고 있으며, 자료의 유지·보수는 275개의 지방 인구등록사무소가 담당하고 있다.

주거에 관한 정보는 주로 건물/주거등록에 포함되어 있다. 건물/주거등록

에는 건물의 주소, 크기, 토지 면적, 용도, 각 주거지의 주거설비, 난방, 소유 등 건물과 주거시설에 관한 구체적인 정보가 있다. 건물/주거등록의 개별 건물과 주거지 관련 정보는 각 지자체가 수시로 갱신하며, 이를 중앙정부의 주택부(Ministry of Housing and Urban Affairs)가 종합하여 총체적으로 관리한다.

사업장(workplace)에 관한 정보는 사업등록(The Business Register)에 있다. 사업등록은 1970년대 초에 세금원천징수를 위해 만들어진 다양한 행정등록체계를 바탕으로 국세청(tax administration)과 중앙세관(central customs)이 관리·운영한다. 그러나 세금징수 목적의 행정등록은 기업(법인)을 중심으로 구성되어 있어 기업에 속한 각 사업장별 자료를 확보할 수 없다. 이런 이유에서 덴마크 통계청은 각 사업장별로 자료를 관리한다. 한편 통계청은 사업장에 자체에 관한 정보는 자체적으로 관리하고 있다. 1970년 세금 원천징수사업체에 관해서는 국세청 등록자료를, 자영업자에 관해서는 관리하고 있다. 사업체 등록자료에는 자영업자와 사업자별 정보로 구성되어 있으며, 사업체의 개인에 대한 고용/피고용 상태, 상시/임시 고용형태, 직업, 산업, 소득 등이 포함되어 있다.

이렇듯 다양한 행정자료를 함께 사용하기 위해서는 이들을 서로 연계할 수 있도록 체계화된 등록번호체계가 필요하다. 먼저 전 국민을 식별하기 위해서는 우리의 주민등록번호와 비슷한 개인별 고유번호가 있어야 한다. 덴마크에서는 개인별 고유번호를 중앙인구등록번호(Central Population Register Number)라 부른다. 이 번호는 생년월일과 함께 무작위로 주어지는 번호, 논리적으로 번호를 점검할 수 있는 점검번호(check number) 등 11개의 숫자로 구성 되어있다. 예를 들어 <표 4-4>의 중앙인구등록번호는 1975년 9월 23일에 태어난 여자의 것이다.

<표 4-4> 덴마크의 중앙인구등록번호 체계

번호	태어난 날짜	태어난 달	태어난 해	태어난 세기	무작위 번호	점검번호 (& 성별)	총 자릿수
자릿수	2	2	2	1	2	1	10

출처 : 덴마크 통계청 내부 자료

전 국민의 주거지와 국가의 모든 건물들을 식별하기 위해서 건물 및 주거별 번호가 있어야 한다. 일반적으로 주소는 건물과 주거지의 식별을 위한 것이다. 그러나 주소는 개인의 신상정보를 보여줄 뿐 아니라 통계적으로 처리하기도 불편하다. 이런 이유에서 중앙인구등록이나 건물/주거등록 등의 등록 자료는 부화된 주소번호를 사용하고 있다. 주소번호는 지역번호 3자리, 도로번호 4자리, 건물번호 3자리, 건물 입구 1자리, 층 2자리, 주거입구 4자리 등 16자리의 체계적인 정보로 구성되어 있다(<표 4-5> 참조).

<표 4-5 > 덴마크의 주소번호 체계

내역	지역 Municipality Code	도로 Road Code	건물 House Number	건물입구 House Letter	층 Floor	주거입구 Side/Door Number	총 자릿수
자릿수	3	4	3	1	2	4	17

출처 : 덴마크 통계청 내부 자료

또한 일에 관한 정보를 처리하기 위해서는 일하는 곳, 즉 사업체에 대한 식별번호가 있어야 한다. 행정적으로 사업체는 국세청이 관리하고 있다. 사업체는 민간기업, 중앙정부, 지방정부로 분류되며 각각이 고유한 번호체계를 갖고 있다. 통계청은 이를 통합하여 상응하는 통계청 자신의 고유번호체계를 만들어 관리한다.

덴마크 통계청은 개인, 주소, 사업체의 번호체계를 바탕으로 덴마크 통계청은 여러 종류의 행정자료를 연계하여 항목별 통계를 만든다. 예컨대 중앙인구등록번호로 국가 전체의 인구통계를, 주소로 행정구역별 인구통계를 산

출한다. 또한 중앙인구등록의 주소를 바탕으로 가구를 구성하여 가구관련통계를, 가족관계 자료를 바탕으로 가족을 구성하여 가족관련 통계를 산출한다. 특히 정확한 가구구성을 위해 개인별 주소의 정확성을 확인하고 오류를 수정하는 과정에서 건물/주거등록자료를 사용한다. 고용관련 통계는 개인의 중앙인구등록번호와 사업체등록에 속한 개인의 중앙인구등록번호를 연계하여 산출한다.

(2) 행정자료 관리제도의 변천

덴마크는 1981년에 센서스를 전통적 전수조사에서 행정자료로 완전히 대체함으로써 세계에서 가장 먼저 행정자료 기반의 센서스를 실시한 국가가 되었다. 또한 인구센서스뿐만 아니라 모든 국가통계를 행정자료로 이용하여 생산하는 국가이기도 하다. 덴마크는 오랜 센서스의 역사를 가지고 있다. 덴마크는 1743년에 징병의 목적으로 출생과 사망 통계를 작성하였고, 1769년 이래 지속적으로 센서스가 실시해 왔다(<표 4-6> 참조).

<표 4-6> 덴마크의 센서스 관련 주요 사건 연혁

연도	사건
1743	출생과 사망 기록(징병 목적)
1769	센서스 실시
1924	지역인구등록(Local Population Registers)
1963	중앙인구등록(Central Population Register) 위원회
1968	중앙인구등록(Central Population Register)
1970	세금원천징수(Tax at Source)
1975	소득 통계
1976	행정등록기반 센서스(Register-based Census) 착수
1978	건물/주거등록(Building and Dwelling Register) 포함
1981	총체적 행정등록기반 센서스(Register-based Census) 실시

출처 : Borchsenius(2000)

센서스에서 행정자료를 활용하는 기반은 1924년에 시행된 지방별 지방인구등록(Local Population Register)사업으로 마련되기 시작하였다. 특히 1968년에 갖추어진 중앙인구등록(Central Population Register) 체제는 행정자료 활용의 중요한 계기가 되었다. 중앙인구등록은 인구의 이중 등록을 방지하고 세금 공제 도입에 따른 세제 개혁에서 전국적으로 인구를 중앙에서 관리하는 사업이었다. 이때 지방인구등록을 중앙에서 통합하기 위해 중앙인구등록의 개인번호와 주소지 번호의 도입하였다. 이 사업은 개인 및 주소지 번호체계를 갖추어 행정자료를 행정전산망을 통해 활용할 수 있는 기반을 마련하였다는 점에서 중요하다(Borchsenius, 2000). 더 나아가 중앙인구등록은 다른 행정체계의 기준으로 활용되어 덴마크는 행정자료를 통합하는 과정에서 상이한 기준체계 때문에 발생하는 혼란을 겪지 않게 되었다.

덴마크는 1970년까지 전통적 방식으로 센서스를 수행하였다. 그렇지만 이미 1960년부터 인구관련의 국적 항목과 주택관련의 편의시설 및 주거기간 항목에서 행정자료를 센서스에 활용하기 시작하였다. 특히 1969년에 중앙인구등록으로 인구통계를 시험적으로 산출하였다. 그 결과는 그다지 만족스럽지 않았으나, 자료의 오류뿐만 아니라 행정등록을 담당하는 기관과 통계청 사이의 협력관계에서 문제점들을 발견하는 계기가 되었다. 이러한 문제점들을 보완하여 1970년에 실시한 센서스 결과와 비교하여, 등록자료를 점검하였으며, 그해 말에는 중앙인구등록자료만을 활용하여 각 지역별, 성별, 연령별, 혼인상태별 인구통계를 산출·발표하였다. 이와 더불어 그해 센서스에서는 건물관련의 주거단위의 수 항목, 경제활동관련의 직업 및 사업체 항목으로 행정자료를 활용하기 시작하였다. 1976년에 건물과 주거를 제외한 센서스 항목들에 대하여 행정등록 기반의 센서스를 실시하였다. 1978년에는 행정자료 기반 센서스를 실시하는데 가장 어려운 부분에 하나인 건물/주거등록(Building and Dwelling Register)을 체계화하였다. 그 결과, 1981년에는 세계 최초로 행정자료 기반 센서스를 실시하게 되었다(<표 4-7> 참조).

<표 4-7> 덴마크의 항목별 행정자료 대체 연도

시작 연도	항목	시작 연도	항목	시작 연도	항목
인구관련		1970	사업체	1960	편의시설
1976	성별	1976	사회경제지위	1960	거주기간
1970	혼인상태	1976	소득	건물관련	
-	모국어	교육관련		1981	건축기간
1960	국적	1976	교육정도	-	건축재
-	종교	가구관련		1981	건물의 주용도
경제관련		1976	가구종류/인원	1970	주거단위 수
1976	경제활동상태	1976	가족종류/인원	-	난방시스템
1976	종사상 지위	주택관련		-	건물의 좌표
1976	산업	-	주택 크기	-	복도면적
1970	직업	1976	방의 수	-	건물의 층수

- : 원래 공표하지 않는 항목

출처 : Laihonon(1999)에서 재구성

이러한 역사적 배경과 함께 덴마크는 행정자료기반 통계작성을 위해 다음과 같은 제도적 기반을 마련해왔다. 덴마크가 행정자료기반 통계작성에 성공할 수 있었던 데는 제도적, 특히 법적 제도의 밑받침이 크다. 대표적인 법은 1966년에 제정된 덴마크 통계법(Act on Statistics Denmark)이다. 이 법은 명시적으로 덴마크 통계청의 독립을 보장했다는 점에서 큰 의의가 있다. 아울러 모든 통계는 통계청에서 중앙관리 하도록 법제화하였다. 그러므로 정부의 각 부처가 통계자료를 생산하기 위해서는 통계청에 의뢰하여야 한다. 또한 모든 개인, 기업, 단체, 정부부처는 통계청에서 요구하는 정보를 제출할 의무를 가지고 있다.

이밖에도 행정자료에 관한 각종 등록법은 행정자료기반 센서스의 성공적인 정착에 중요한 역할을 하였다. 주요 관련법으로는 1968년 제정된 인구등록법(Act on Population Registration), 1975년에 제정된 사업체등록법(Business Register Act), 1977년에 제정된 건물및주거등록법(Act on

Registration of Building and Dwellings) 등이 있다(<표 4-8> 참조).

<표 4-8> 덴마크 통계관련 법

연도	법	주요 내용
1966	덴마크 통계법	통계청의 독립성, 통계의 중앙집중화, 통계청의 자료접근권
1968	인구등록법	인구 등록
1975	사업체등록법	사업체 등록
1977	건물및주거등록법	건물과 주거 등록
1978	공중등록법	자료 반출과 접근 제한

출처 : 덴마크 통계청 내부 자료

이와 더불어 덴마크에는 행정자료기반 센서스 자료를 보호하기 위한 다양한 정책적, 제도적 장치가 있다. 먼저 법적 장치로 1978년에 제정한 공중등록법(Public Registers Act)이 있다. 덴마크는 이 법에 근거하여 자료의 반출과 접근을 엄격히 제한하고 있다. 다음은 조직적 장치로 통계청 내부와 외부에서 각종 위원회를 운영하여 통계의 생산과 자료 활용에 대한 견제와 감시를 한다. 또한 통계 관련 보직을 정기적으로 바꾸어 정보의 독점과 남용을 미연에 방지하고 있다. 이밖에도 기본적인 보안대응체계를 구축하여 건물, 사무실, 통제지역 관리 같은 물리적인 통제와 함께 각종 자료보안, 신분노출 방지, 외부인 접근방지 등 제도적 차원의 자료 통제 장치를 두어 정보를 각종 침해로부터 보호하고 있다.

이러한 법적·제도적 장치는 1960년대부터 통계의 중요성을 이해한 덴마크 지도층과 행정자료 활용의 가능성을 인정한 통계청 지도자들이 있었기에 가능했다. 또한 1960년대에는 개인정보에 대한 국민들의 인식도 민감하지 않았기 때문에 행정등록 기반의 센서스를 도입하기 수월했다. 이런 역사적 맥락 뿐 아니라 덴마크의 사회제도는 정확한 행정자료를 수집하는 기반이 되고 있다. 각종 사회보장제도가 주거지역을 기반으로 운영되기 때문에 중앙인구등록 및 건물/주거 등록에 대한 개인정보가 정확하다. 또한 연금보험 운영과 관련되어 대부분의 취업자들이 직능단체/노조에 가입되어 있어 고용관련

정보를 체계적으로 수집할 수 있다. 이러한 사회제도는 덴마크가 행정자료를 바탕으로 신뢰할 수 있는 센서스를 수행할 수 있는 기반이 되고 있다.

라. 핀란드

(1) 행정자료 기반 센서스 자료 생산 방법

덴마크와 마찬가지로 행정자료를 연결하여 통계를 산출하기 위해서는 개별 행정자료를 서로 연결할 수 있는 번호체계가 필요하다. 그러한 번호체계는 개인과 개인이 속한 가구의 주거지, 개인이 속한 사업체의 번호와 연결하여 주거관련, 경제활동 관련 통계를 산출한다. 따라서 개인별 중앙인구등록번호, 중앙인구등록번호가 포함되는 건물/주거 번호 및 법인/사업체 번호에 대한 체계가 있어야 한다. 중앙인구등록번호는 우리나라의 주민등록번호, 덴마크의 중앙인구등록번호와 매우 유사하게 생년월일, 일련번호, 점검번호로 구성되어 있다(<표 4-9> 참조).

<표 4-9> 핀란드의 중앙인구등록번호 체계

번호	0	5	0	1	5	9	-	5	0	4	2
내역	태어난 날짜		태어난 달		태어난 해		태어난 세기	일련 번호			점검번호

출처 : Myrskylä(2000)

건물/주거번호(Domicile Code) 건물번호와 주거번호 두 가지 종류의 번호로 이루어져 있다(<표 4-10> 참조). 사업체번호는 국세청이 부여하고 있다. 통계청에서 자체적으로 사업체등록번호를 부여하고 있다. 그러한 이유는 국세청은 법인에게만 관리번호를 부여하기 때문에 개별 사업장에 대해서는 통계청이 자체적으로 번호를 부여하여 이용하고 있다.

<표 4-10> 핀란드의 건물/주거 번호 체계

내역	건물번호 Building Code					주거번호 Dwelling Code	
	지역 Municipal	소지역 Village	토지 Real Estate	점검번호 Check Code	건물 Building	입구 Door	주거 Dwelling
자릿수	3	3	4	1	3	1	3

출처 : Myrskylä(2000)

(2) 행정자료 관리제도의 변천

핀란드는 다른 북유럽 국가들과 같이 인구센서스의 오랜 역사를 가지고 있다(<표 4-11> 참조). 스웨덴의 식민지배 기간인 1749년에 스웨덴 정부는 식민통치의 목적으로 최초의 인구센서스를 실시하였다. 러시아 식민지배 기간인 1865년에는 통계국이 개설되었고, 1879년에 통계연감이 발간되었다. 이러한 역사는 핀란드 국민에게 통계의 중요성을 인식하는 바탕이 되었다. 이러한 배경에서 식민지배로부터 벗어난 1948년에 핀란드 정부는 최초로 국민총소득을 집계하였고, 1990년에는 덴마크에 이어 세계에서 두 번째로 행정등록 기반 센서스를 실시하였다.

<표 4-11> 핀란드의 센서스 관련 주요 사건 연혁

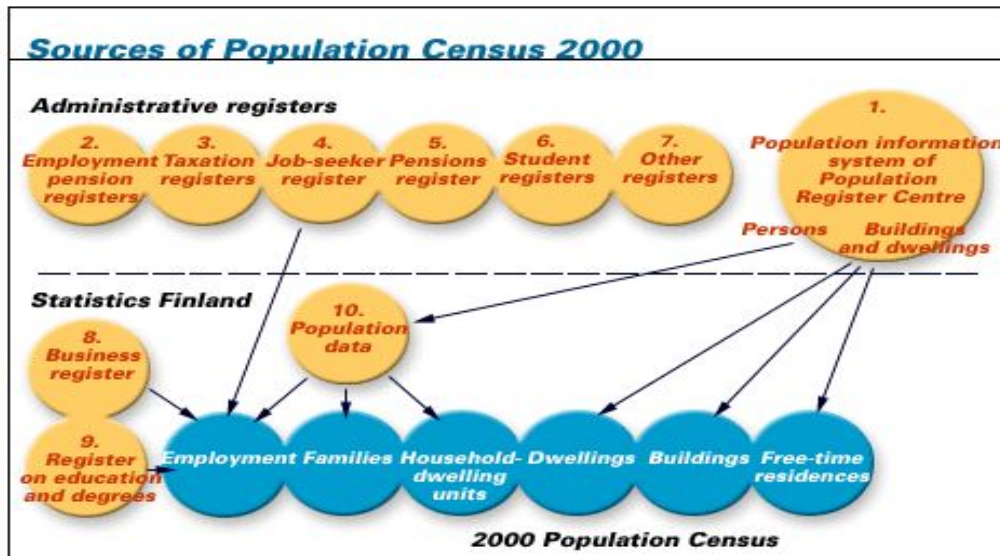
연도	사건
1748	스웨덴 통계 담당 전문기관 설립
1749	최초 핀란드 국민 통계 수집
1865	핀란드 통계사무소(Statistical Office) 설립
1879	최초의 핀란드 통계연감 출판
1948	국민총소득(Gross National Income) 공표 시작
1971	Statistics Finland로 개명
1990	전적 행정자료 기반 센서스(Register-based Population Census)

출처 : 핀란드 통계청 내부 자료

핀란드에서 행정자료가 통계의 목적으로 활용되기 시작한 것은 1960년대 부터이다. 1960년대 말 중앙인구등록(Central Population Register)이 실행되면서 모든 국민에게 우리의 주민등록번호와 비슷한 개인별 번호를 주었다. 이후 번호는 납세나 고용연금보험 등과 같은 일반 행정영역에서도 사용되었다. 이를 바탕으로 1970년 센서스부터 중앙인구등록과 납세등록의 개인자료를 활용하기 시작하였다. 1970년대와 1980년대에 걸쳐 행정자료의 활용은 계속 확대되었다(<표 4-11> 참조). 1985년 센서스에서는 개인의 경제활동 관련 자료들만이 질문지를 통해 조사되었다. 이러한 과정을 거쳐 1990년에는 총체적인 행정등록 기반의 센서스를 실시하였다.

이런 역사적 과정을 거쳐 탄생한 핀란드 행정자료기반 센서스는 현재 다음과 같은 구성 체계를 갖고 있다. 핀란드의 인구센서스는 중앙인구등록, 건물/주거등록, 사업체등록, 세금등록, 연금등록 등 여러 종류의 독립된 행정자료를 기반으로 하고 있다((그림 4-2) 참조). 센서스에서는 여러 종류의 행정자료들을 연결하여 필요한 통계를 산출한다. 예를 들어, 센서스에서 가장 중요한 인구통계는 주로 중앙인구등록자료와 건물/주거등록자료를 바탕으로 산출한다. 따라서 행정등록 가운데 중앙인구등록과 건물/주거등록이 가장 기본이 된다. 인구등록자료는 중앙과 지역사무소의 인구등록센터에서 매년 수정하여 통계청으로 보낸다. 이 자료는 연령, 성별, 결혼상태, 언어, 국적, 종교, 출생국, 가족관계, 주거 등 행정자료기반 센서스에 필수적인 개인정보를 포함하고 있다. 인구통계는 센서스의 가장 기본적인 항목으로 인구등록자료는 다른 자료보다 먼저 1975년 센서스에서 조사를 대체하였다.

건물/주거등록자료 역시 인구등록센터가 관리하고 있다. 건물/주거의 편의 시설이나 건축물 등에 변경이 있으면 각 지역사무소에 신고를 해야 한다. 중앙사무소에서는 이를 취합하여 총체적으로 관리한다. 1980년 센서스에서 질문지 조사를 바탕으로 건물/주거 자료를 수집하고, 이를 바탕으로 등록자료를 검토·수정한 후, 1985년 센서스에서 본격적으로 건물/주거등록자료를 활용한 산출하였다. 더 나아가 1987년 이후에는 매년 건물/주거에 관한 통계자료를 만들게 되었다(Harala and Nieminen, 1998).



(그림 4-2) 핀란드 행정등록 기반 센서스 구성도
(출처 : 핀란드 통계청 내부자료)

경제활동에 관한 통계 산출은 사업체등록에 바탕을 두고 있다. 이외에도 국민의 경제활동 상태를 파악하기 위해서는 중앙인구등록, 건물/주거등록, 교육등록, 국세등록, 연금보험등록 등 다양한 행정자료를 함께 사용한다. 사업체등록자료는 핀란드 통계청에서 자체적으로 관리하고 있으며, 이를 바탕으로 각 사업체에서 연금자료와 연계하여 개인별 자료를 수집하고, 국세등록에서 소득을, 교육등록에서 교육수준 등을 취합한다. 이렇듯 다른 자료에 비해 상대적으로 복잡한 통계산출절차로 말미암아 경제활동에 관한 자료는 1990년에야 비로소 행정등록 기반의 센서스로 바뀌었다(<표 4-12> 참조).

핀란드 행정자료기반 센서스의 특징은 다음과 같다. 첫째, 핀란드 행정자료는 체계적인 주거등록 자료를 기반으로 하고 있다. 여기에는 몇 가지 요소가 중요한 역할을 했다. 무엇보다도 오랜 스웨덴, 러시아 통치시절을 통해 엄격한 주거등록체계가 유지되었다. 이러한 역사적인 전통과 동시에 주거를 기반으로 한 사회적인 혜택을 통해 주거등록 자료의 질을 향상시키고 있다. 이러한 혜택과 동시에 엄격한 법률적인 제재도 실시하고 있다. 예를 들면, 핀란드에선 주거 등록은 일주일 이내에 완료해야 한다. 이를 위반할 경우 엄격한 법률적인 제재가 따른다.

<표 4-12> 핀란드의 항목별 행정자료 대체 연도

시작 연도	항목	시작 연도	항목	시작 연도	항목
	인구관련	1990	사업체	1985	편의시설
1975	성별	1990	사회경제지위	1985	거주기간
1975	혼인상태	1975	소득	건물관련	
1975	모국어	교육관련		1985	건축기간
1975	국적	1980	교육정도	1985	건축재
1970	종교	가구관련		1985	건물의 주용도
경제관련		1975	가구종류/인원	1985	주거단위의 수
1985	경제활동상태	1975	가족종류/인원	1985	난방시스템
1985	종사상 지위	주택관련		1975	건물의 좌표
1990	산업	1985	주택 크기	1985	복도면적
1990	직업	1985	방의 수	1985	건물의 층수

출처 : Laihonon(1999)에서 재구성

둘째, 행정자료를 기반으로 하는 센서스를 통해 비용과 효율적인 측면에서 성공적인 결과를 산출하고 있다. 각종 통계자료의 생산은 10배 증가했다. 동시에 비용은 1/4로 감소시켰다.

셋째, 핀란드 통계청에서는 자료유출이나 개인 사생활 보호를 위해 다음과 같은 조치를 취하고 있다. 기본적으로 자료에 대한 접근을 제한한다. 대부분 사람은 허락된 범위 내에서만 자료에 접근할 수 있다. 따라서 각 부문의 전문가들은 부분적으로만 자료에 접근한다. 동시에 통합자료는 외부유출이 엄격하게 금지된다. 정보기관을 비롯한 여타 어떠한 국가기관의 접근도 금지되고 있다. 물론 외부 유출 시에도 개인식별번호를 삭제하는 등 각종 제약을 부가하고 있다. 핀란드에서 정보보안을 위하여 제정한 각종법안이 <표 4-13>에 나와 있다.

<표 4-13> 핀란드의 정보보호 관련 법

연도	관련 법
1994	통계법(The Statistics Act)
1995	EU의 자료보호 강령(Data Protection Directive of the EU)
1997	지역통계 규정(Council Regulation on Community Statistics)
1999	정부행위공개법(Act on the Openness of Government Activities)
1999	개인자료법(Personal Data Act)

출처 : 핀란드 통계청 내부 자료

마. 덴마크와 핀란드에 대한 평가

덴마크와 핀란드의 센서스 방식은 행정자료를 활용하여 전통적 센서스의 여러 문제들을 해결할 수 있는 하나의 방안을 제시하였다. 국민들의 응답 부담은 사라진 반면, 과거 질문지를 사용해 조사되었던 거의 모든 항목들에 대해서 행정자료에 의한 통계를 산출하게 되었다. 자료의 질에서도 질문지 조사에 비해 나아진 것으로 평가되고 있다(Nordberg et al., 2001).

행정자료의 활용은 전통적 방식의 센서스가 가지고 있던 비용과 응답부담의 문제점을 해결했을 뿐만 아니라 전수에 바탕을 둔 자료를 연도별로, 지역별로 생산해 낼 수 있게 되었다. 매년 다양한 변수에 대한 시계열 자료의 생산은 새로운 연구 가능성을 제시하였다. 소지역 추정 가능성이 가능해지고, 변동에 관한 통계 산출이 손쉬워졌다(Myrskylä 1999). 이와 함께, 종단적 자료(longitudinal data)의 생산도 가능해져 직접적으로 인과성을 살펴볼 수 있게 되었다.

행정자료를 활용한 센서스 방식이 많은 장점을 가지고 있지만, 다른 한편으로는 이 방식이 갖고 있는 문제점들도 간과해서는 안 된다. 행정자료에 오류가 있는 경우 그 영향이 전통적인 질문지 조사 방식보다도 더 큰 영향을 미칠 수 있다. 예컨대 큰 사업장의 주소에서 지자체가 잘못 기록된 경우 사업장에 소속되어 있는 모든 종업원들이 실제와는 다른 지자체로 연결된다. 이 경우 행정 자료에 있는 오류들이 지역 통계에 상당히 큰 오차를 가져 올

수 있다.

행정자료가 다양한 부처에서 생산되는 경우 동일한 항목을 중복하여 조사하는 경우가 많은데, 이 때 개념 규정이 약간씩 차이를 갖게 되는 경우 이 차이를 조화시켜서 새로운 변수를 만들어야 한다(Spieker, 1999). 또한 법령 개정이나 정책의 변화에 따라 수집되는 자료에 관해 변화가 발생할 경우 통계의 연속성을 확보하기 어렵다는 문제가 있다.

제3절. 우리나라 행정자료 기반 센서스의 가능성

1. 이론적 모형

서로 다른 행정자료를 결합하기 위해서는 행정에서 사용하는 일반 문건이나 서식에 서로를 연결할 수 있는 연계키(matching key)가 필요하다. 이를 위해서 먼저 행정의 대상이 되는 행위의 주체(actor) 모두에게 행위자 개체를 식별할 수 있도록 개체별로 고유한 코드를 부여하여 한다. 이와 함께 이렇게 부여한 코드를 모든 행정에서 공통으로 사용하여야 한다. 행위의 주체는 개인과 법인이 있다. 먼저 개인의 경우를 살펴보면, 이름은 개인을 식별할 수 있는 코드이다. 그렇지만 같은 이름을 가진 사람이 있기 때문에 이름만으로는 고유한 코드가 되기에 부족하다. 이런 문제를 해결하기 위해 만든 코드가 주민등록번호다. 주민등록번호는 원칙적으로 모든 국민이 고유한 코드를 하나씩 갖고 있다. 또한, 1962년부터 40여 년 동안 모든 행정에서 사용해 온 탓에 행정자료의 연계키의 역할을 하기에 충분하다.

법인의 경우는 국세청에서 사용하는 사업자등록번호가 있다. 이 번호는 주민등록번호처럼 모든 법인을 식별할 수 있는 고유한 코드이다. 주민등록번호와 차이는 이 번호를 국세청 이외에도 모든 행정기관에서 공통으로 사용하지 않는다는 것이다. 또 다른 문제는 법인체는 여러 사업체나 사업장의 집합체인 경우가 있다는 점이다. 예를 들어 전국 규모의 법인체로부터 세금을 걷기 위해서는 하나의 법인으로 등록하여도 관계가 없지만, 지역마다 있는 서비스센터나 유통망의 지방세, 건물의 방재, 범죄 발생, 근로감독을 위해서

는 지역별로 동사무소와 구청에 개별적으로 등록하여야 한다. 이럴 경우 법
인체를 식별하는 고유 코드는 지역별 사업체를 식별할 수 있으면서 동시에
모법인 소속임을 알 수 있도록 만들어져야 한다. 그래야만 각각의 행정 서류
를 연결하여 국세, 지방세, 지역 행정 등에 두루 사용할 수 있기 때문이다.

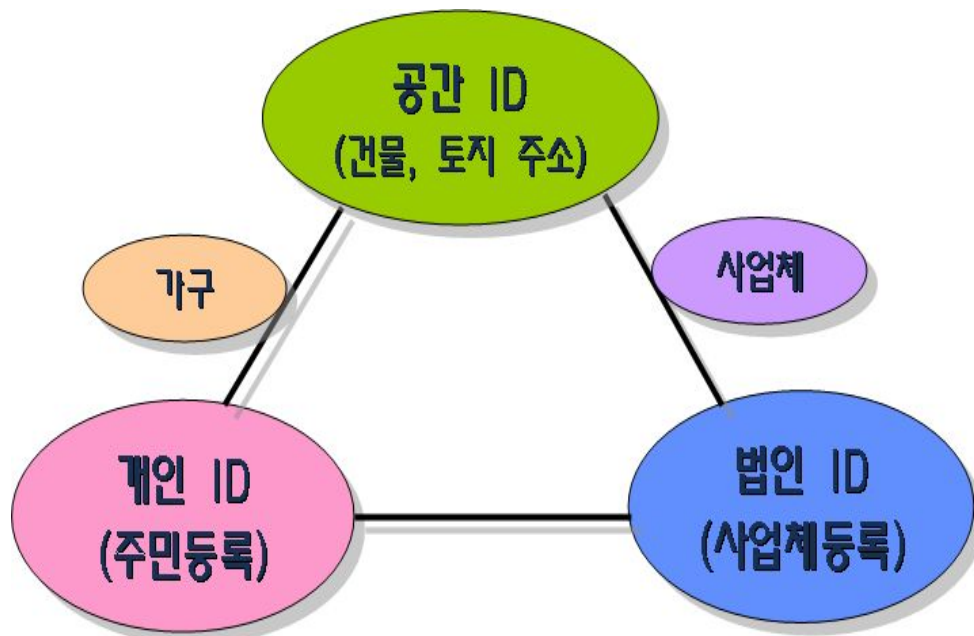
그렇지만 사업체 수준의 법인에 대한 고유 코드를 우리는 아직 가지고 있
지 않다. 앞으로 국세, 지방세, 산업, 노동, 복지 등에서 두루 사용할 수 있도
록 사업체 수준의 법인 코드 체계를 만드는 것이 아마도 행정자료 결합을
위한 첫 걸음이 될 것이다. 코드 개발에서 코드를 부여할 사업장 및 사업체
의 수준과 관할지역의 범위를 정해야 하며, 코드에서 적어도 모법인, 관할지
역, 산업 분류 등은 식별이 가능하도록 해야 공통으로 사용할 수 있을 것이
다.

여기서 우리는 행정은 권역별로 나뉘어 관리되기 때문에 행정대상의 행위
가 발생하는 공간적 위치(spatial location)를 파악할 수 있어야 행정적으로
처리 할 수 있다는 점을 알게 된다. 앞서 예로 든 전국 규모의 법인체의 경
우 개별 서비스센터의 관할 동사무소, 구청, 시청을 알기 위해 위치를 파악
해야 한다. 공간적 위치를 파악하는 데는 여러 가지 방법이 있지만 일반적으
로 문서에서는 주소를 사용한다. 이때 주소는 개인과 법인별로 고유한 주소
를 가져야 한다. 다시 말해 개인과 법인은 반드시 하나의 주소와 상응하여야
한다. 그렇지 않으면 한 개인이나 법인을 관할하는 동사무소나 구청, 또는
도청 및 시청이 여럿이 될 수 있기 때문이다.

그렇지만 현재의 주소인 지번체계로는 이렇게 할 수 없다. 농촌경제 시대
에 만들어진 필지 중심의 지번체계로는 건물과 건물 안에 있는 거처 및 사
무공간을 식별하기 어렵다. 현재 진행 중인 새주소 체계는 건물을 식별할 수
있도록 만들었기 때문에 지번체계보다는 고유한 코드에 한 걸음 접근해 있
다고 할 수 있다. 하지만 이 또한 건물 내부의 거처와 사업장을 식별하는 체
계까지 갖추지는 않았다. 행정자료 결합을 위해 가장 먼저 해야 할 일이 모
든 거처와 사업장을 식별할 수 있는 주소체계를 만들어 행정기관에서 사용
하는 것이 될 것이다.

이제 행정자료를 결합하기 위해서는 기본적으로 개인별, 사업체 수준의

법인별, 이들이 속한 공간별 위치를 식별할 수 있는 고유한 코드가 있어야 하며, 이 코드를 모든 행정에서 공통으로 사용해야 한다는 점을 알게 되었다. (그림 4-3)은 이를 도형화 한 것이다. 이 그림은 공간에 대한 고유 코드(공간 ID)의 중요성을 한층 더 부각시킨다. 행정자료만으로 같은 공간 ID를 가진 개인의 고유번호(개인 ID)를 모으면 가구(household)가 된다. 또한 같은 공간 ID를 가지고 또 같은 법인 ID를 가진 사업장들을 묶어 사업체(establishment)로 정의할 수 있다. 다시 말해 행정자료만을 결합하여 가구와 사업체를 식별하여 행정이 수월하면서도 효율적일 수 있도록 할 것이다.



(그림 4-3) 행정자료 연계를 위한 행정자료의 기본틀
(출처 : 덴마크 통계청 내부자료)

2. 인구주택총조사에서 행정자료의 활용가능성

이제 이런 모형에 비추어 우리나라의 행정자료를 살펴보면, 개인 ID로는 주민등록, 법인 ID로는 사업체등록, 공간 ID로는 주소가 있다. 다시 말해, 행정자료를 이용할 기반이 갖추어진 것이다. 이에 착안하여, 김설희·강계화(2005)는 인구주택총조사의 조사항목을 각종 행정자료를 결합하여 조사 없이

통계자료를 생산 할 수 있을 것이라고 제시하였다. 그들이 제시한 항목별 행정자료의 활용가능성은 <표 4-14>와 같다.

<표 4-14> 2005년 센서스 항목의 행정자료 활용가능성

항목	활용가능성 (%)	항목	활용가능성 (%)	항목	활용가능성 (%)
가구원에 관한사항		경제활동상태	50	거주 층	30
성명	100	종사상 지위	0	자동차대수	100
성별	100	산업	50	주차시설	0
나이	70	직업	0	난방시설	0
가구주와 관계	100	근로장소	50	점유형태	0
남북이산가족	70	혼인상태	80	입차료	0
교육정도	70	혼인 년 월	80	주인가구/소유	80
종교	0	총 출생아 수	80	주택에 관한 사항	
아동보육실태	0	추가계획자녀 수	0	거처종류 및 층수	80
활동제약	70	고령자 생활비	0	연건평	80
5년 전 거주지	80	가구에 관한사항		대지면적	80
통근통학 여부	50	가구구분	100	총방수	80
통근통학 장소	50	거주기간	100	건축년도	100
이용교통수단	0	사용 방수	40	편익시설수	80
통근통학시간	0	주거시설형태	20	전체단순평균	
				51.6	

출처 : 김설희·강계화(2005)

이때 활용되는 행정자료는 주민등록초본, 호적등초본, 이산가족 방북신청자 현황, 교육행정자료시스템, 장애인 증명서, 근로소득원천징수영수증, 건축물대장총괄표제부, 등기부 등본, 자동차등록증 등이 있다. 아래 (그림 4-4) 주민등록, (그림 4-5)는 건축물대장의 항목과 인구주택총조사의 항목별 연결을 예시로 제시한 것이다(김설희·강계화, 2005).

이러한 예시적 제시는 마치 주민등록과 건축물대장에서 지금이라도 센서스의 통계의 일부 항목을 생산할 수 있을 것처럼 보여준다. 그렇지만, 이들의 이러한 시도는 이론적 수준에서 1 대 1 대응의 가능성을 제시한 것일 뿐이라는 데 현실적인 어려움이 있다(이명진외, 2006). 주민등록과 인구주택 총조사 항목의 1 대 1 대응은 주민등록이 실제 생활을 그대로 반영하는 것을 전제로 한다. 특히 주민등록은 세대주를 중심으로 세대별로 등재되기 때문에 세대와 상응하는 가구와 실질적으로 일치해야 한다. 개념적으로 보면, 세대는 주거 및 생계를 같이하는 사람들의 집단이며, 가구는 “1인 또는 2인 이상이 모여서 취사, 취침 등 생계를 같이하는 생활단위”(통계청, 2005: 10)로 매우 비슷하다. 그렇지만 주민등록의 세대(de jur)는 아파트 청약이나 채권채무 관계 등 행정적인 이해관계와 밀접하게 관련이 있기 때문에 실제 주거관계(de facto)와 다른 경우가 많다. 이런 경우 인구주택총조사의 가구주 및 가구원과 주민등록의 세대주 및 세대원이 달라 주민등록으로 인구센서스의 개인 수준 자료를 생산할 경우 오류가 발생한다.

또한 단독가구가 증가하는 추세에서 세대는 가구와 일치하지 않을 가능성이 높다. 예를 들어 원룸에 사는 30대 초반의 미혼여성은 주민등록을 부모에게 두고 있을 가능성이 높다. 이럴 경우 그 여성은 인구주택총조사에서는 단독가구로 조사되지만 주민등록에서는 부모와 함께 사는 세대로 남아있게 된다. 또 단독주택에 여러 세대가 사는 경우와 단독주택에 부모와 분가한 아들 가족이 함께 사는 경우도 주민등록으로 구별하기가 쉽지 않다.

이런 문제를 해결하기 위해서는 행정자료로 가구를 정의할 수 있는 개념 정립이 필요하며, 이때 가구는 실제 생활단위를 가장 잘 반영할 수 있어야 한다(Borchsenius, 2000). 보통은 거처를 중심으로 생활단위를 파악하여 거처(주거) 가구(dwelling household)란 개념을 쓴다.²³⁾ 이럴 경우 개별 거처를 식별할 수 있는 주소체계가 중요한 역할을 한다. 스웨덴과 노르웨이는 거처별 주소체계가 잘 정비되지 않아 행정자료 기반의 인구센서스로 전환하는

23) 가구에 대한 상세한 개념은 The Canberra Group(2001)과 The Task Force on Statistics for Rural Development and Agricultural Household Income(2005)을 참조할 것. City Group에 관한 사항은 UN Statistics Division의 홈페이지인 <http://unstats.un.org/unsd/methods/citygroup/index.htm>를 참조할 것.

데 애를 먹었다. 행정에서 1947년부터 개인식별번호를 사용했음에도 불구하고 거처별 등록(dwelling register)이 되어있지 않아 2000년 인구센서스에서 가구에 관한 자료는 전통적인 조사방식으로 확보할 수밖에 없었다(Statistics Sweden, 1999). 스웨덴은 2005년 인구센서스를 완전히 행정자료를 기반으로 시행하고자 다양한 연구와 시험조사를 진행하였다(ONS, 2003; 홍두승외, 2002).

우리나라에서도 이견외(2004)와 이명진외(2006)는 불완전한 주소체계로 말미암아 가구별로 건축물대장과 조사자료를 연계하기 매우 어렵다는 사실을 발견하였다. 우리의 현 주소체계로는 거처는 말할 것도 없고 건물을 식별하기조차 어렵다. 물론 새 주소체계가 도입이 완료되는 2011년이면 이 문제는 조금 나아질 것이다. 그러나 행정자료를 기반으로 통계를 생산하기 위해서는 여전히 거처별로 식별할 수 있는 주소체계가 필요하며, 이에 대한 준비도 필요하다.

이런 문제들로 행정자료와 인구주택총조사 사이에는 많은 차이가 나타날 수 있다. 다음 장에서는 인구주택총조사와 주민등록의 차이를 구체적으로 살펴보고 그러한 차이가 얼마나 나는지를 파악한다. 이를 바탕으로 주민등록으로 인구주택총조사를 과연 대체할 수 있는지, 대체는 얼마나 가능한지, 만약 차이가 난다면 어떤 패턴이 있는지, 그러한 패턴으로 둘 사이의 조정이 가능한지 등을 살펴보고자 한다. 그리고 그 다음 장에서는 건축물 관련 행정자료를 총조사에 실제로 활용할 수 있는지, 활용하는 데 문제점은 무엇이 있는지를 살펴본다.

제5장. 인구주택총조사자료와 주민등록자료 비교

제1절. 서론

주민등록으로 인구주택총조사를 대체하더라도 지금까지 축적해온 인구센서스 자료와 시계열을 맞추기 위한 조정 필요하다. 이러한 조정에는 많은 시간이 소요된다. 이것은 두 자료의 비교·대조가 그다지 쉽지 않을뿐더러, 인구센서스가 5년에 한번 있는 조사이기 때문이다. 조정을 위한 조사 설계가 인구주택총조사에 포함되어야 하며, 그 조사결과를 바탕으로 조정 작업을 진행해야 한다. 따라서 2 차례의 시험조사만 거쳐도 10년이 소요된다. 이미 행정자료 기반으로 인구센서스를 집계하는 나라에서 새로운 방식의 도입에 20년 이상의 기간이 소요된 것을 보면 그러한 조정 작업이 얼마나 긴 기간을 준비해야 하는 일인지를 알 수 있다.

이런 문제의식에서 이 장에서는 인구주택총조사와 주민등록을 비교하여 주민등록으로 인구주택총조사를 대체할 가능성과 대체 시 발생할 문제점을 살펴본다. 두 자료 사이에 비교할 내역들은 매우 다양하여 항목 차이, 개념 차이, 조사방법이나 집계방법의 차이 등 다뤄야 할 범위는 무척 넓다. 그렇지만 이 연구는 시작 단계에 있는 것이므로 행정적으로 가장 활용이 많이 되는 읍면동 수준의 집계인구와 가구 및 세대를 비교한다.

읍면동 수준의 집계인구를 비교하는 것은 두 자료 사이의 차이를 살펴보는 것 이외에도 여러 가지 이점이 있다. 행정에서 사용되는 인구와 실제 거주 인구 사이의 차이를 대략적이거나 알 수 있어 행정에서 실제 인구가 필요한 경우에는 참고할 수 있다. 또한 만약에 센서스를 행정자료로 대체할 경우에도 시계열을 조정하기 위해서는 가장 먼저 해야 할 작업 가운데 하나가 읍면동 수준의 인구를 조정하는 것이다.

인구 다음으로 총조사의 중요한 통계가 가구통계다. 따라서 여기서는 총조사의 가구와 주민등록의 상응하는 단위인 세대를 비교한다. 또한 이동이 많은 젊은 층에서는 주민등록지와 실제 생활하는 장소가 다른 경우가 많을

수 있다. 이런측면에서 연령별 인구도 비교한다. 이와 함께, 2004년 인구주택 총조사 시험조사에서 획득한 두 지역의 시험조사 자료와 주민등록 자료를 개인별 수준에서 매칭하여 상호호응의 정도를 살펴본다.

제2절. 인구 비교²⁴⁾

1. 연구배경

인구주택총조사의 집계는 16개 시도, 234개 시군구, 3,572개 읍면동 수준으로 되어 있으며, 지역별·연령별 인구는 읍면동 수준까지 공표하고 있다. 우리나라에서 집계수준에서 주민등록과 인구주택총조사를 다룬 연구는 그다지 많지 않으며, 그러한 연구도 시도 수준에서 다룬 것이 대부분이다. 그러한 연구로는 한국인구학회(1998)과 홍두승외(2002)가 있는데 이들의 경우 두 자료의 개괄적인 차이를 보여주는 데에 목적을 두어 두 자료에서 나타나는 차이를 상세히 이해하기 어렵다.

좀 더 상세한 연구는 읍면동 수준의 집계자료를 이용한 김근영외(2003)의 연구가 있다. 그들은 인구주택총조사에서 생산되는 통근·통학인구를 바탕으로 1995년과 2000년도 사이의 통근통학인구를 계산하기 위해 주민등록인구를 사용하였다. 먼저 수도권에 거주하는 인구 및 가구(세대)를 대상으로 1995년과 2000년 인구주택총조사와 주민등록 자료를 비교하였다. 수도권의 읍면동은 1995년에는 1,131개, 2000년에는 1,140개였다. 읍면동별로 주민등록 인구와 인구주택총조사인구의 차이의 평균은 1995년에 505.1명, 2000년에는 755.3명으로 2000년이 더 많은 차이를 보였다. 또한 가구에 대해서도 같은 방식으로 계산하여 평균차이는 1995년에 735.9가구, 2000년에 793.5가구로 역시 2000년에 더 큰 차이를 나타냈다. 이 차이를 바탕으로 그들은 주민등록 대비 인구주택총조사 인구의 비율의 차이를 평균하여 매년 주민등록 대비 인구주택총조사 인구의 비율을 구하는 보정방식(adjustment method)을 제시하였다.

24) 이진(2007)을 바탕으로 작성

그러나 그들의 보정방식은 매우 단순하여 실제 현상을 얼마나 잘 반영할 지 의문이다. 현실은 매우 복잡하여 두 자료 사이의 차이가 평균 차이에서 3σ 이상인 경우도 1995년에 15, 2000년에 18 읍면동이 있었으며, 2σ 이상인 경우도 각각 11, 30 읍면동이 있었다. 2σ 이상인 경우를 살펴보면, 인구수의 표준편차는 1995년에 1,374.7명, 2000년에 933.1명이었으니, 주민등록과 인구주택총조사의 인구수 차이는 평균차이 + 2σ 로 1995년 3,254.5명 이상(평균 + 2σ), 2000년에는 2,621.5명이나 된다. 주민등록으로 보았을 때 수도권 읍면동의 평균 인구수는 약 2만 명으로 이정도 차이는 10% 이상이 된다. 다시 말해, 주민등록 대비 인구주택총조사 인구의 차이가 10% 이상인 읍면동이 1995년에는 26개(2.3%) 이상, 2000년에는 48개(4.2%) 이상 있다.

이렇듯 큰 차이는 일괄적인 계량적 보정으로만 해결할 수 있는 것이 아니다. 개별 읍면동의 특성을 살펴보고 어떤 이유에서 주민등록과 인구주택총조사의 차이가 나는지를 살펴 그와 관련된 변수의 통제가 필요하다. 이런 이유에서 여기서는 읍면동별로 집계인구수의 차이를 살펴보고, 지역의 특성과 어떤 연관이 있는지를 찾아본다.

2. 자료

인구주택총조사와 주민등록이 얼마나 서로 잘 호응하는지를 살펴보기 위해서는 인구주택총조사의 개인 및 가구별 자료와 주민등록을 1 대 1로 비교하는 것이 가장 좋은 방법이다. 그러나 일반 연구자가 인구주택총조사와 주민등록의 개인별 자료를 확보하기란 현실적으로 불가능하다.²⁵⁾ 이런 제약으로 말미암아 이 연구는 통계청이 집계하여 홈페이지(<http://www.nso.go.kr>)에 공표하는 읍면동 수준의 인구집계자료를 이용하였다.

인구주택총조사는 2005년 11월 1일부터 15일까지 진행되었으며, 11월 1일 0시를 기준으로 인구수를 파악하여 읍면동별로 공표하고 있다. 한편 행정자치부는 매년 12월 31일 24시를 기준으로 주민등록에 등재된 인구수를 읍면

25) 인구주택 총조사와 주민등록 자료를 개인 수준에서 비교한 경우는 통계청 내부 연구(이 내성, 2006)가 있다.

동별로 집계하여 공표하고 있다. 한 가지 유의할 점은 두 자료가 모두 정확하더라도 기준 시점이 약 2달 차이가 있어 두 집계자료 사이에는 어느 정도 차이가 있을 수밖에 없다는 사실이다. 그 사이에 출생한 사람, 사망한 사람, 장기체류를 위해 해외로 떠난 사람, 이민을 간 사람, 해외에서 살다 돌아온 사람이 있으며, 다른 지역으로 이사를 간 사람도 있을 것이다. 이사한 사람만을 놓고 보면, 우리나라에서 1년에 이동하는 인구가 약 20% 정도니 2달 동안에는 약 3%가 이동한다고 할 수 있다. 그러한 이동이 읍면동 안에서 이루어지는 경우도 있기 때문에 이동으로 인해 나타날 수 있는 두 집계자료의 차이는 많아야 3% 정도일 것이다. 이 연구에서는 그러한 차이를 구체적으로 파악할 수 없기 때문에 다루지 않기로 한다. 이 연구의 목적이 주민등록으로 인구주택총조사를 대체할 수 있는지 살펴보는 것으로, 이 연구의 관심이 큰 오차에 있기 때문에 이 정도의 오차는 큰 문제가 되지 않을 것으로 본다.

3. 비교 결과

가. 전국 수준

국가 전체로 보면 2005년 인구주택총조사의 집계인구수는 47,041,434명이며, 주민등록의 집계인구수는 48,782,274명으로 주민등록인구수가 1,740,840명 더 많았다. 이를 주민등록인구수 대비 인구주택 총조사 인구의 비율로 보면 96.4%다. 이러한 패턴은 1995년과 2000년 인구주택총조사에서도 비슷하게 나타났다(<표 5-1>참조). 1995년에는 주민등록인구가 45,981,910명으로 총조사 인구 44,608,726명보다 1,373,184명 많았으며, 그 비율이 96.9%였다. 2000년에는 주민등록인구가 47,976,730명으로 총조사인구 46,136,101명보다 1,840,629명 많았으며, 그 비율이 96.0%였다. 이렇듯 주민등록인구가 인구주택총조사인구보다 많은 이유는 조금 다른 인구집계방식에서 발생한다. 주민등록인구에는 인구주택 총조사에서는 제외되는 해외연수·유학생, 해외취업자 및 파견근로자 등이 포함되어 있다.

이외에도 오류로 인한 차이가 있을 수 있다. 주민등록이 없거나 중복으로

등록된 경우도 있고, 인구주택 총조사에서 누락되거나 중복 조사된 경우도 두 집계자료의 차이로 나타난다. 주민등록의 중복이나 누락은 행정상의 오류나 제도의 남용으로 발생하기 때문에 알기도 쉽지 않고 추정하기도 쉽지 않다. 특히 주민등록의 직권말소가 채권채무관계 및 재판절차의 진행을 위해 이용되는 경우에 주민등록 누락으로 이어지기 쉽다.

인구주택 총조사의 누락과 중복은 사후조사를 통해 밝혀낸다. 2005년 인구주택 총조사의 순누락률은 0.9%로 약 42만 6천명이 누락된 것으로 추정되었다(이지연, 2007a). 그 이전 총조사를 살펴보면 순누락률은 1985년에 0.82%, 1990년 -0.04%, 1995년에 1.25%, 2000년에 1.56%로 대체로 1% 수준이다(김형석, 1999).²⁶⁾ 이러한 검증결과는 인구주택총조사의 집계인구수는 실제 거주인구수에 매우 가깝다는 사실과 함께 인구주택총조사의 안정성과 정확성을 보여준다.

나. 시도 수준

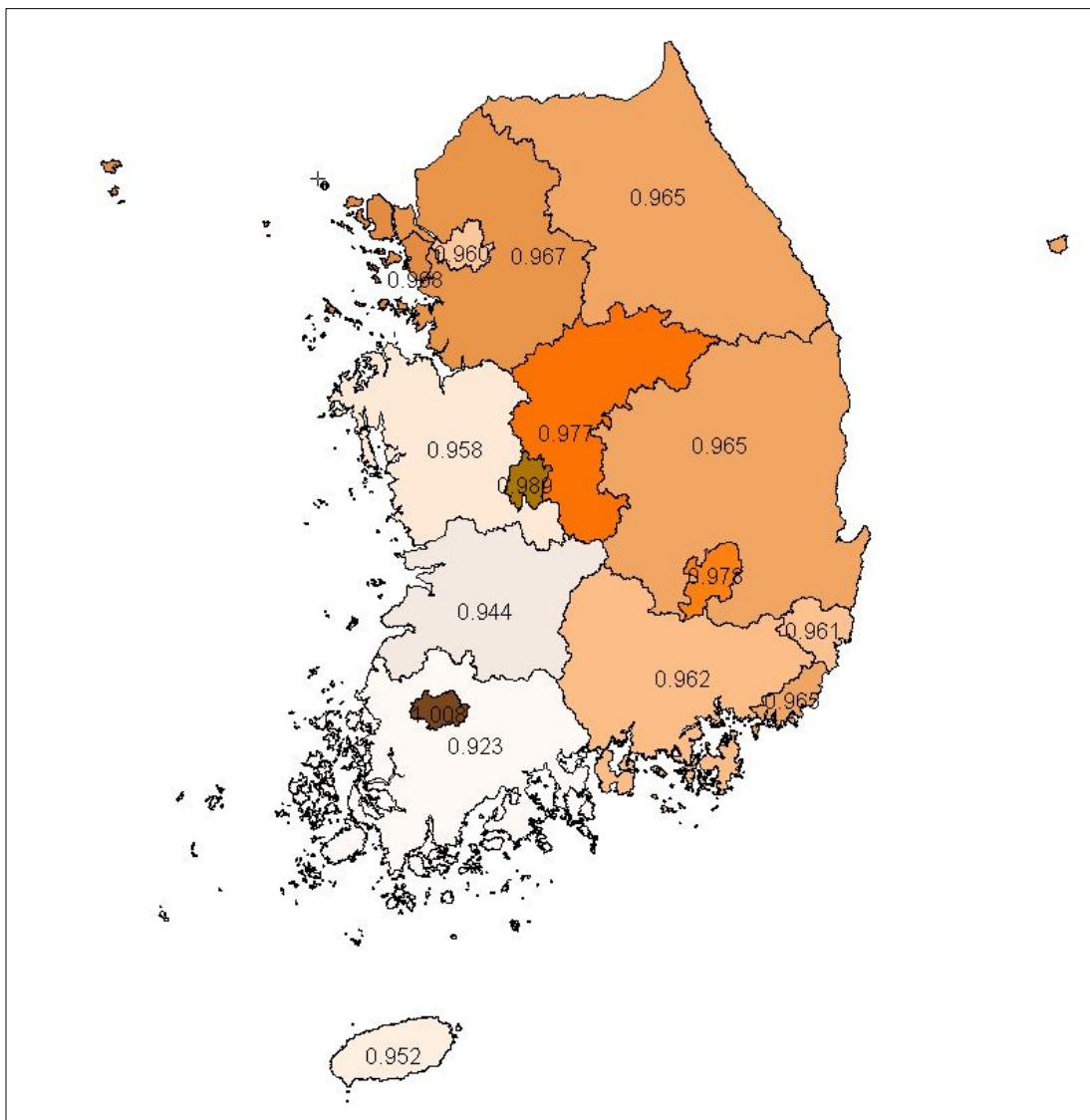
16개 시도별로 주민등록 대비 인구주택총조사 인구의 비율을 살펴보았다((그림 5-1)). 먼저 이들의 평균은 96.5%로 국가 전체의 96.4%와 큰 차이가 없다. 또한 표준편차는 1.9%로 많은 시도가 96% 근처에 있어 대체로 안정된 분포를 보이는 편이다. 주민등록 대비 인구주택 총조사 인구 비율이 가장 낮은 지역은 전라남도로 92.3%, 그 다음이 전라북도로 94.4%여서 호남지역에서 주민등록과 총조사의 차이 많이 나타났다. 또한 제주도는 95.2%, 충청남도는 95.8%로 대체로 우리나라의 남서지역에서 주민등록인구보다 실제 거주인구가 적은 것으로 나타났다.

시도수준에서 과거 인구주택총조사와 비교해보았다(<표 5-1>). 이 표에서는 주민등록인구 대비 주민등록인구와 총조사인구의 차이를 계산하여 이의 백분율을 차이율이라고 제시하였다. 지난 10년 동안의 변화를 보면, 약간의 변동은 있으나 대체로 변동의 폭이 크지 않다. 여전히 전남의 차이율이 가장

26) 이 정도의 누락률은 일본의 0.5% 보다는 못하지만 미국의 1.5%와 비슷하며 대체로 유럽보다는 나은 편이다.

크며, 전북, 충남 순을 유지하고 있다. 제주도는 10년 전에 비해 차이율이 증가하였으며, 새로 광역시로 편입된 울산의 차이율은 증가하는 모습을 보인다. 반면 광주의 차이율은 오히려 줄었다. 전남의 경우 지난 10년 동안 차이율이 증가하였으며, 차이율 또한 다른 곳에 비해 큰 편이다. 이런 경우는 개인 수준의 자료를 직접 대조하여 그 원인을 찾아 볼 필요가 있을 것이다.

(그림 5-1) 시도별 주민등록인구 대비 인구주택 총조사인구의 비율



<표 5-1> 주민등록인구와 총조사인구의 시도별 차이, 1995, 2000, 2005년

	1995년			2000년			2005년		
	총조사 인구	주민 등록 인구	차이율 (%)	총조사 인구	주민 등록 인구	차이율 (%)	총조사 인구	주민 등록 인구	차이율 (%)
전국	44,608,726	45,981,910	3.1	46,136,101	47,976,730	4.0	47,041,434	48,782,274	3.6
서울	10,231,217	10,595,943	3.6	9,895,217	10,373,234	4.8	9,762,546	10,167,344	4.0
부산	3,814,325	3,892,972	2.1	3,662,884	3,812,392	4.1	3,512,547	3,638,293	3.5
대구	2,449,420	2,485,977	1.5	2,480,578	2,538,212	2.3	2,456,016	2,511,306	2.2
인천	2,308,188	2,362,132	2.3	2,475,139	2,562,321	3.5	2,517,680	2,600,495	3.2
광주	1,257,636	1,287,134	2.3	1,352,797	1,375,212	1.7	1,413,644	1,401,745	(0.8)**
대전	1,272,121	1,268,432	(0.3)**	1,368,207	1,390,510	1.6	1,438,551	1,454,638	1.1
울산	-	-	-	1,014,428	1,044,161	2.9	1,044,934	1,087,648	3.9
경기	7,649,741	7,811,468	2.1	8,984,134	9,280,013	3.3	10,341,006	10,697,215	3.3
강원	1,466,238	1,530,000	4.3	1,487,011	1,559,042	4.8	1,460,770	1,513,110	3.5
충북	1,396,728	1,442,191	3.3	1,466,567	1,504,722	2.6	1,453,872	1,488,803	2.3
충남	1,766,854	1,855,346	5.0	1,845,321	1,930,234	4.6	1,879,417	1,962,646	4.2
전북	1,902,044	2,009,651	5.7	1,890,669	2,006,500	6.1	1,778,879	1,885,335	5.6
전남	2,066,842	2,186,808	5.8	1,996,456	2,134,629	6.9	1,815,174	1,967,205	7.7
경북	2,676,312	2,775,922	3.7	2,724,931	2,813,551	3.4	2,594,719	2,688,491	3.5
경남	3,845,622	3,958,540	2.9	2,978,502	3,108,674	4.4	3,040,993	3,160,431	3.8
제주	505,438	519,394	2.8	513,260	543,323	5.9	530,686	557,569	4.8

* 차이율 = {(주민등록인구-총조사인구)/주민등록인구} * 100

** ()는 총조사인구가 주민등록인구보다 많은 경우

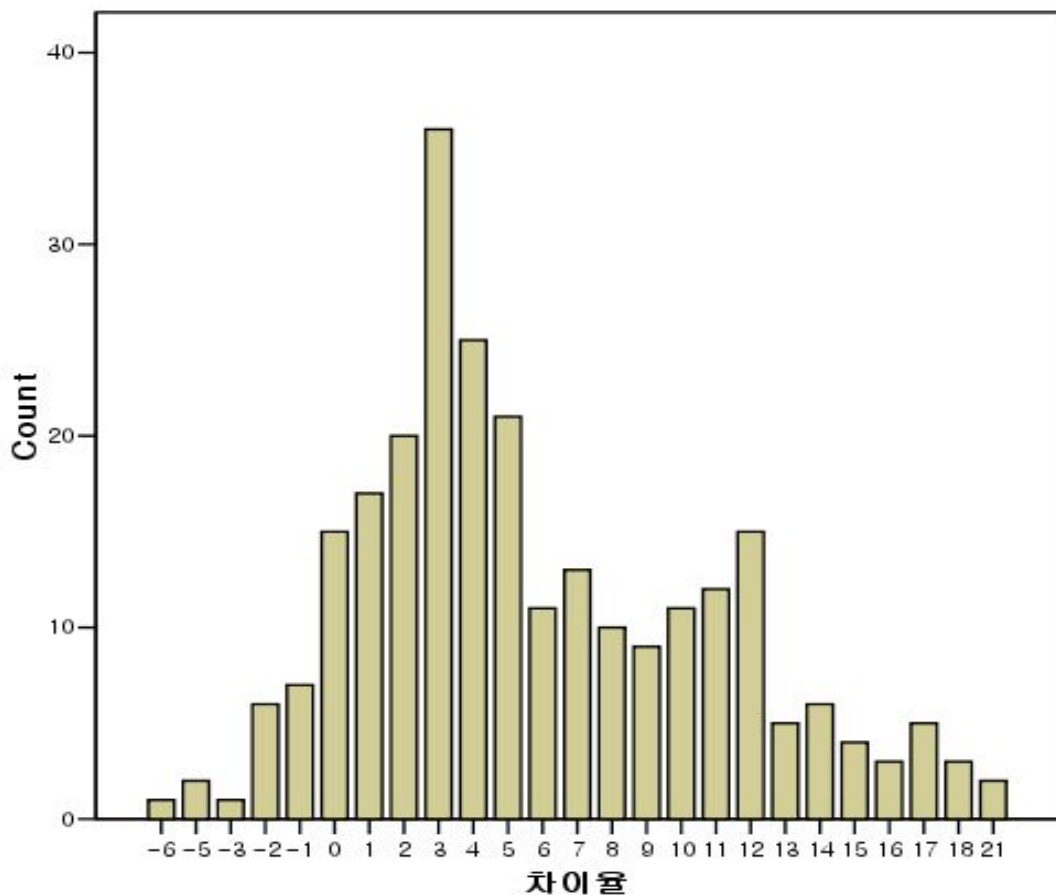
자료출처 : 통계청(각년도)

다. 시군구 수준

전국의 시군구는 234개로, 주민등록 대비 인구주택총조사 인구 비율의 평균은 93.8%, 표준편차는 5.2%, 중간값(median)이 95.1%다. 시도수준에 비해 평균이 2.7%p 감소하고 편차는 크게 늘었다. 아래 (그림 5-2)는 이들의 분포를 2% 간격으로 제시한 것이다. 20개 시군구에서는 총조사인구가 더 많고,

나머지 214 시군구에서는 주민등록인구가 더 많다. 총조사인구가 더 많은 시군구는 편차가 그다지 크지 않다. 반면 주민등록인구가 많은 경우에는 편차가 큰 시군구도 많아 주의를 요한다. 차이율이 3% 이상인 시군구는 159개로 전체 시군구의 67.9%이며, 5% 이상인 시군구는 115개로 49.1%에 이른다. 특히 차이율이 10% 이상인 시군구도 62개로 26.5%나 된다. 이것은 시군구 수준에서 주민등록자료를 인구센서스로 이용하려면 많은 보정 작업이 필요하다는 것을 의미한다.

(그림 5-2) 시군구 수준의 주민등록인구와 총조사인구 차이율 분포



시군구별 차이율을 상자도표로 표시해보았다(그림 5-3). 도가 광역시에 비해 분산이 크다. 차이율의 중앙값도 대체로 도가 광역시보다 큰 편이다. 이

어떤 시군구의 인구는 적게 집계될 것이다.

<표 5-2> 주민등록과 인구주택총조사 인구의 차이가 큰 시군구

	순번	시/도	시/군/구	총조사 인구수 (명)	주민등록 인구수 (명)	차이율 (%)
총조사 인구가 많은 시군구 (3% 이상)	1	경상북도	경산시	240,371	227,766	5.5
	2	강원도	태백시	55,241	52,463	5.3
	3	광주광역시	동구	119,783	114,480	4.6
총조사인구가 적은 시군구 (15% 이상)	1	인천광역시	옹진군	12,261	15,609	21.4
	2	전라북도	임실군	25,682	32,511	21.0
	3	경상남도	합천군	47,651	58,162	18.1
	4	전라북도	순창군	26,250	32,012	18.0
	5	전라북도	진안군	23,915	29,021	17.6
	6	전라남도	신안군	38,429	46,451	17.3
	7	제주도	북제주군	80,419	97,202	17.3
	8	전라북도	장수군	20,486	24,755	17.2
	9	제주도	남제주군	61,421	73,841	16.8
	10	충청북도	괴산군	32,208	38,595	16.5
	11	전라남도	구례군	24,699	29,577	16.5
	12	경상북도	성주군	39,348	46,960	16.2
	13	경상북도	청도군	39,435	46,748	15.6
	14	전라남도	장성군	41,682	49,165	15.2
	15	전라북도	부안군	55,240	65,018	15.0
	16	부산광역시	강서구	45,726	53,808	15.0

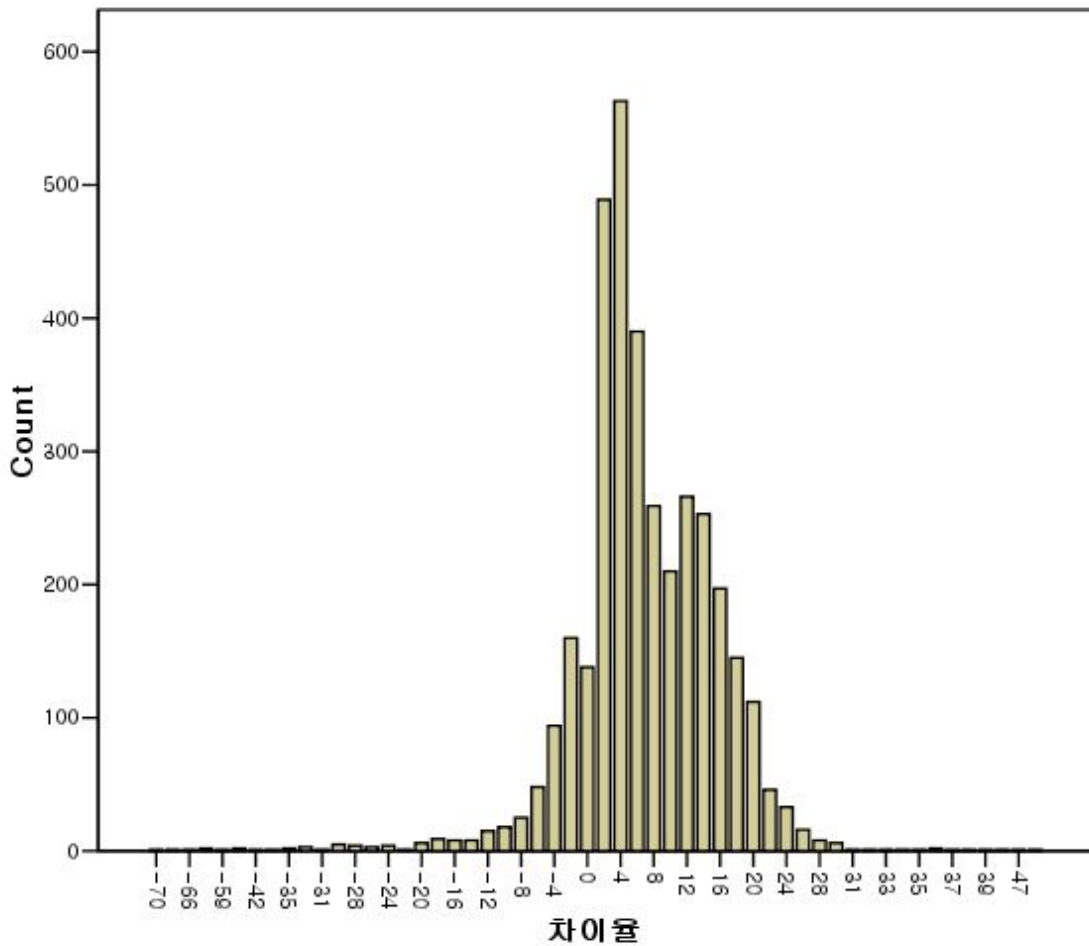
라. 읍면동 수준

전국에는 총 3,584개의 읍면동이 있다. 그 가운데 몇몇 읍면동은 주민등록과 인구주택총조사를 한데 묶어 집계를 낸다. 이런 이유에서 자료에 나타난 읍면동의 수는 3,572개다. 먼저 주민등록 대비 인구주택총조사의 비율은 평균이 93.7%, 표준편차가 12.0%, 중간값이 94.8%다. 시군구 수준과 비교하면 평균과 중간값은 비슷하고 표준편차는 2배가량 증가하였다.

(그림 5-4)는 주민등록 대비 인구주택총조사 인구비율의 분포를, 차이율이

30% 이상인 경우를 제외하고, 2% 간격으로 나타낸 것이다. 3,082(86.3%) 읍면동에서는 주민등록인구가 많고, 나머지 489개(13.7%) 읍면동에서는 주민등록인구가 적어 주민등록인구가 많은 경향을 반영하고 있다. 전체적으로는 시군구의 분포와 흡사한 모습을 보인다. 특히 95% - 100% 사이에 가장 집중되어 있는 모습과 85% - 90% 사이에 또 하나의 약한 집중현상이 있는 모습이 비슷하다.

(그림 5-4) 주민등록 세대와 총조사 가구의 읍면동별 차이율 분포



<표 5-3>은 작은 차이와 큰 차이의 분포를 대조하여 살펴보기 위해 중심부는 촘촘히, 주변부는 느슨하게 빈도를 제시한 표이다. 주민등록 대비 인구

주택총조사의 인구차이 비율(차이율)²⁷⁾이 $\pm 3\%$ 이내인 읍면동(표의 어두운 부분)은 모두 953개로 전체 읍면동의 26.7%이며, $\pm 5\%$ 이내인 읍면동은 1,597개로 전체의 44.7%이고, $\pm 10\%$ 이내의 읍면동은 2,333개로 65.3%에 이른다. 나머지 1,239개(34.7%) 읍면동은 차이율이 10%보다 크다. 시군구 수준에서는 26.5%이던 비율에 비해 8%p 가량 증가하였다. 또한 차이율이 15%보다 큰 읍면동도 574개로 16.0%나 되어 적지 않으며, 30% 이상 차이가 나는 경우도 37개소로 1% 가량을 차지하고 있다.

<표 5-3> 읍면동 수준의 총조사인구/주민등록인구 분포

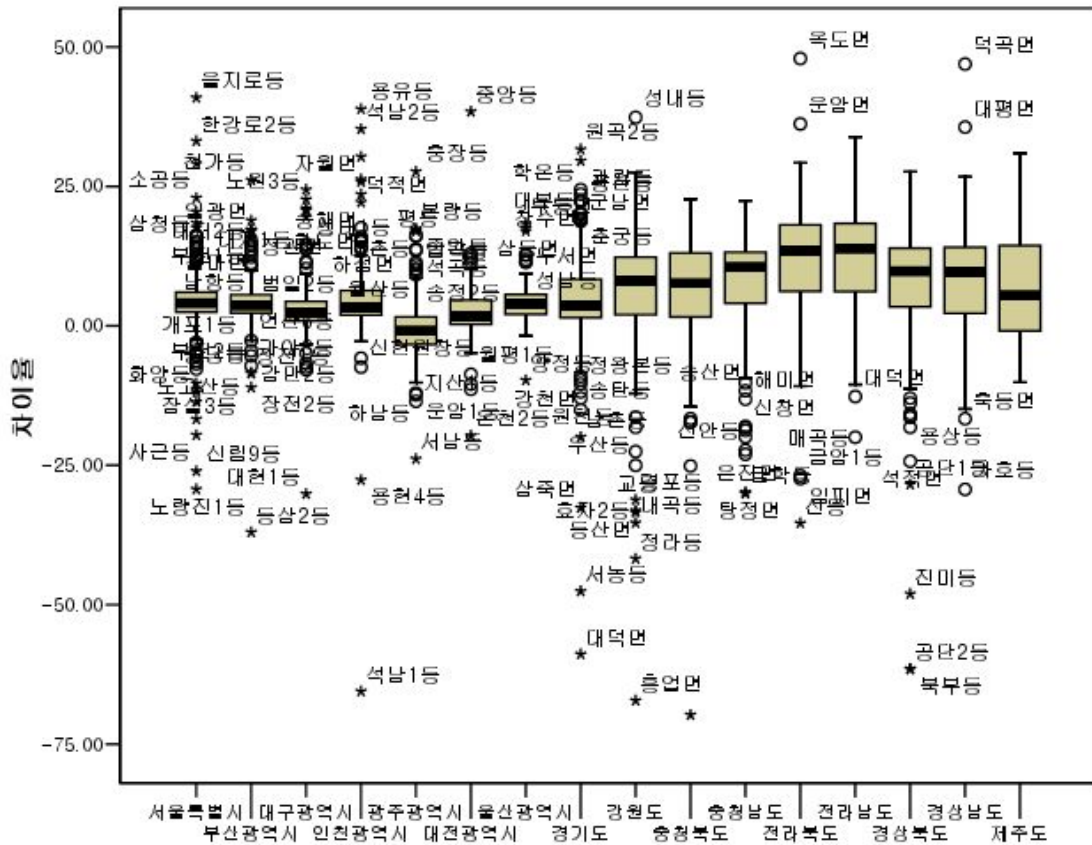
총조사인구/주민등록인구 비율(%)	빈도	%
85 미만	519	14.6
85-90	632	17.7
90-95	675	18.9
95-97	551	15.4
97-99	520	14.6
99-100	185	5.2
100-101	122	3.4
101-103	126	3.5
103-105	93	2.6
105-110	61	1.7
110-115	33	0.9
115 초과	55	1.5
전체	3,572	100

시도, 시군구 수준과 비교하면, 전체적으로 소지역 수준으로 갈수록 주민등록과 인구주택총조사의 차이가 커진다. 이러한 결과는 큰 지역수준에서 지역 내의 오차들이 서로 상쇄되기 때문에 나타나는 현상이다. 한 예로 서울의 용산구에 살고 사는 사람이 자녀의 취학을 위해 서초구에 주민등록을 두고 있는 경우를 들 수 있다. 이럴 경우 시도 수준에서는 주민등록과 인구주택총조사의 인구집계에는 차이가 없지만, 시군구 수준과 읍면동 수준에서는 차이가 발생한다.

27) 차이율은 100-총조사인구/주민등록인구 비율(%) 이다.

읍면동별 총조사와 주민등록 차이의 분포를 상자도표로 그려보았다((그림 5-5)). 시군구 경우와 마찬가지로 광역시에 비해 도의 분산이 컸다. 이것은 광역시보다 도의 읍면동 수준 주민등록인구가 작아 조그만 인구의 차이도 비율로는 크게 나타나는 현상일 수도 있기 때문에 상세히 살펴보아야 할 부분이다. 전체적으로 차이율의 분산은 시군구에 비해 커졌다. 이 또한 주민등록인구가 애초에 작아서 나타나는 현상일 수도 있다. 이상값은 시군구와 달리 양과 음으로 고르게 나타났으며, 주민등록인구가 적은 경우도 많아졌고, 그 차이도 매우 커졌다.

(그림 5-5) 주민등록 세대와 총조사 가구의 시군구별 차이율 상자도표



이렇듯 전반적으로 소지역 수준으로 갈수록 분산이 커지고 차이도 커질뿐더러, 큰 차이를 보이는 읍면동이 많아 그러한 차이가 나는 이유를 살펴볼

필요가 있다. 이럴 경우 차이가 큰 지역부터 살펴보는 것이 적절하여, 차이율이 30% 이상인 지역을 <표 5-4>에 제시하였다. 표의 위쪽에는 총조사인구가 많은 20개 읍면동을, 아래쪽에는 총조사인구가 적은 16개 읍면동을 제시하였다.

가장 눈에 띄는 지역은 서울시 송파구 잠실1동과 잠실2동으로 다른 지역과는 비교도 할 수 없을 만큼 큰 차이(500% 이상)가 난다. 이러한 차이는 이 지역이 재개발 중이어서 주민들의 주거상태가 불안정하기 때문에 생긴 것으로 특이한 경우(outlier)라 할 수 있다. 또한 서울시 성북구 월곡3동은 차이율이 66.8%며 인구 또한 매우 적은데 이 지역 역시 재개발이 진행 중이어서 이런 차이가 발생한 것으로 보인다. 이들 만큼 차이가 크지는 않지만 경기도 성남시 분당구 판교동도 차이율이 36.6%이고 인구도 많지 않은 것도 재개발 때문이다. 이와 같은 지역들에서 나타나는 차이는 인구집계에 주민등록을 이용할 경우 재개발 지역에 대해서는 특별한 조치를 취해야 한다는 점을 시사한다. 그렇지만 재개발 지역에서는 주거인구 자체가 불안정하기 때문에 인구집계가 큰 의미를 갖지 않을 수도 있다. 따라서 재개발 지역에 대해서 정확한 인구집계를 위해 어떠한 조치가 필요한지에 관해서는 앞으로 많은 논의가 있어야 할 것이다.

경상북도 경산시는 대학이 많아 학생들이 주민등록을 본가에 두고 학교 근처에서 생활하기 때문에 총조사인구가 주민등록인구보다 많은 것으로 보인다. 또한 경상북도 구미시는 산업단지로 주민등록을 본가에 두고 구미시에서 직장생활을 하는 사람, 특히 미혼남녀나 기혼이지만 떨어져 혼자 사는 경우가 많아 총조사인구가 주민등록인구보다 많을 것이다.

<표 5-4> 주민등록 대 인구주택총조사 인구의 차이가 30% 이상인 읍면동

	순번	시도	시군구	읍면동	총조사 인구(명)	주민등록 인구(명)	차이율 (%)
총조사인구가 많은 경우	1	서울특별시	송파구	잠실2동	15	0	*
	2	서울특별시	송파구	잠실1동	152	26	584.6
	3	충청북도	충주시	달천동	8,699	5,124	69.8
	4	강원도	원주시	홍업면	10,703	6,402	67.2
	5	인천광역시	서구	석남1동	23,031	13,913	65.5
	6	경상북도	구미시	공단2동	7,255	4,491	61.5
	7	경상북도	경산시	북부동	18,811	11,646	61.5
	8	경기도	안성시	대덕면	14,175	8,922	58.9
	9	경상북도	구미시	진미동	22,954	15,504	48.1
	10	경기도용인	기흥구	서농동	14,204	9,625	47.6
	11	강원도	삼척시	정라동	12,102	8,537	41.8
	12	부산광역시	영도구	동삼2동	6,812	4,972	37.0
	13	전라북도	익산시	신동	24,478	18,079	35.4
	14	강원도	강릉시	내곡동	14,827	10,955	35.3
	15	강원도	춘천시	동산면	2,270	1,701	33.5
	16	강원도	춘천시	효자2동	19,332	14,522	33.1
	17	경기도	안성시	삼죽면	4,671	3,522	32.6
	18	강원도	강릉시	경포동	7,275	5,545	31.2
	19	대구광역시	북구	대현1동	11,091	8,522	30.1
	20	충청남도	아산시	탕정면	17,034	13,093	30.1
총조사인구가 적은 경우	1	서울특별시	성북구	월곡3동	233	702	66.8
	2	전라북도	군산시	옥도면	2,289	4,395	47.9
	3	경상남도	합천군	덕곡면	909	1,713	46.9
	4	서울특별시	중구	을지로동	1,052	1,780	40.9
	5	인천광역시	중구	용유동	2,176	3,559	38.9
	6	대전광역시	동구	중앙동	2,031	3,297	38.4
	7	강원도	삼척시	성내동	7,997	12,769	37.4
	8	경기도 성남시	분당구	판교동	572	902	36.6
	9	전라북도	임실군	운암면	1,366	2,141	36.2
	10	경상남도	진주시	대평면	853	1,325	35.6
	11	인천광역시	서구	석남2동	13,211	20,410	35.3
	12	전라남도	여수시	묘도동	971	1,468	33.9
	13	서울특별시	용산구	한강로2동	2,723	4,073	33.1
	14	경기도 안산시	단원구	원곡2동	5,833	8,531	31.6
	15	제주도	북제주군	우도면	1,241	1,796	30.9
	16	인천광역시	옹진군	영흥면	2,778	3,986	30.3

그 외의 지역에 대해서는 눈에 띄이게 뚜렷한 패턴을 찾기가 어렵다. 이것은 실제로 특성이 없어서기보다는 개별 지역에 대해 세세히 알지 못하기 때문이다. 따라서 주민등록으로 인구집계를 하기 위해서는 지역의 특성에 대해 잘 알아야 하며, 특히 그러한 특성 가운데 인구집계에 영향을 주는 변수를 찾을 수 있어야 한다. 앞으로 주민등록을 기반으로 인구집계를 할 경우 그러한 변수들은 주민등록인구를 보정하는 주요 수단이 될 것이다. 또한 지역에 관한 지식은 보정치를 산출하기 위해 표본지역을 선정할 때도 도움을 줄 것이다. 이런 측면에서 주민등록으로 인구집계를 하려면 지역에 관한 연구와 전문지식이 기반이 될 것이기 때문에 이에 대한 지식기반 확충이 필수적이다.

● 예시 : 서울의 경우

조금 구체적인 지역의 특성을 살펴보기 위해 읍면동 수준에서 서울의 주민등록 대비 인구주택총조사 인구의 비율을 지도에 표시하였다((그림 5-6)). 지도에서는 색깔이 진할수록 두 집계자료의 차이가 큰 읍면동이다. <표 5-4>에서 30% 이상의 차이율을 보인 송파구 잠실1, 2동, 성북구 월곡3동, 용산구 한강로2동, 중구 을지로동은 모두 가장 짙은 색깔로 표시되어 있다. 지도는 행정동 이름과 함께 지역의 특성을 파악하는 데 도움이 된다. 표에서는 위치를 파악하기 어려웠던 용산구 한강로2동의 경우 지도에서는 삼각지와 용산역 사이에서 전자상가 지역임을 쉽게 알 수 있었다. 이 지역에서도 역시 용산역 주변 재개발이 활발하기 때문에 두 집계자료에 큰 차이가 난 것으로 보인다. 또한 중구 을지로동은 청계천 북단으로 집계자료의 차이는 청계천 복원과 관계가 있을 것으로 보인다.

차이율이 30%보다는 작지만 10% 이상인 지역을 살펴보면 강북 도시에 상당히 많이 분포하고 있다. 그 중 가장 눈에 띄이는 지역이 종로구다. 종로구의 대부분은 차이율이 15%에서 30% 사이에 있다. 특히 종로1, 2, 3가 지역에서는 차이율이 29% 정도로 나타났다. 상업지구로 유동인구가 많고, 실거주자의 파악이 수월하지 않기 때문인 것으로 보인다. 한편 용산구 대부분

지역과 은평구 진관내동과 진관외동도 큰 차이를 보이는데 이 또한 재개발 때문인 것으로 짐작할 수 있다. 이외에도 마포구 상암동, 서초구 내곡동, 강남구 압구정동, 강동구 강일동 등에서 차이율이 15% 정도로 높다. 이런 지역의 상대적으로 높은 차이율을 이해하기 위해서는 그들 지역에 대한 좀 더 상세한 연구가 필요할 것이다.

(그림 5-6) 서울의 주민등록 대비 인구주택총조사 인구비율의 분포



4. 소결 및 논의

이 연구는 주민등록으로 인구를 집계할 경우 어떤 문제가 있는지를 찾아 보기 위한 탐색적 연구다. 이를 위해 이 연구에서는 읍면동 수준에서 집계한 주민등록과 인구주택총조사의 인구를 비교하였다. 이 연구를 통해 알게 된 가장 큰 사실은 주민등록으로 인구주택총조사의 인구통계를 만들기가 매우

어려우며, 두 자료의 비교 또한 매우 복잡하다는 것이다. 따라서 앞으로 주민등록으로 인구통계를 집계하기 위해서는 상당히 많은 연구와 이를 위한 시험조사 및 병행조사의 과정을 거쳐야 할 것으로 보인다.

두 자료를 비교한 결과, 집계수준이 작을수록 두 자료의 차이는 크고, 차이가 나는 지역의 수도 많다는 사실을 발견하였다. 또한 이에 대해 적절하게 처리를 하지 않으면 주민등록으로 인구를 집계할 경우 지역별 인구에 대한 잘못된 통계를 생산할 가능성이 높다는 사실도 알게 됐다. 이러한 위험을 줄이기 위해 앞으로 인구주택총조사에서는 주민등록과 맞추어 가기 위한 연구와 함께 다양한 조정 작업이 필요할 것이다.

이 연구는 탐색적 연구로 주민등록으로 인구를 집계하기 위해 필요한 연구들을 제안하는 것도 목표 가운데 하나다. 먼저, 주민등록자료로 인구센서스의 인구집계가 가능한지를 좀 더 정확하게 파악하기 위해서는 인구주택총조사의 기준시점과 같은 기준시점에서 집계한 주민등록자료를 이용한 연구가 있어야 할 것이다. 둘째, 주민등록에 포함되어 있는 해외거주자, 이민자, 유학생 등에 관한 자료를 확보하여 이들을 셋째, 주민등록으로 총조사 인구를 파악하기 위해서는 지역의 특성에 따른 보정 작업이 있어야 한다. 여기에는 지역사회에 관한 지식기반확충이 매우 중요하다. 우리나라에서 지역사회에 대한 지식기반이 매우 취약하다. 지방자치제가 시행된 지 10여 년이 지났지만 아직도 중앙집권적 정치와 행정으로 지역에 대한 관심이 적기 때문이다. 앞으로 지역사회 관련 지식기반 확충에는 많은 시간과 노력이 들 것이다.

제3절. 가구와 세대 비교

1. 가구와 세대의 개념 비교

인구주택총조사와 주민등록을 비교하는 작업은 여러 측면에서 진행될 수 있다. 첫 번째, 총조사의 인구수와 주민등록인수를 비교하는 것이다. 두 번째, 총조사의 가구수와 주민등록세대수를 비교하는 것이다. 마지막으로, 총

조사의 남녀 혹은 세대별 집계수와 주민등록 성별 혹은 연령별 집계를 비교하는 것이다. 여기에서는 가구수와 세대수를 비교함으로써 주민등록 자료를 어떻게 총조사에서 활용할 것인가 하는 문제에 함의를 찾고자 한다. 이러한 작업은 가구나 세대가 일상생활에서 중요한 단위로 기능을 하고, 정책적으로 중요한 의미를 갖고 있다는 점에서 의의가 있다고 생각한다.

이를 위해서는 먼저 기초 개념인 세대와 가구를 비교할 필요가 있다. 무엇보다도 이러한 과정이 전제되어야만 주민등록자료와 총조사의 매칭 가능성을 살펴볼 수가 있다. <표 5-5>는 주민등록자료의 세대 개념과 총조사의 가구 개념을 비교하고 있다. 일반적으로 사전에서 세대(世帶)는 현실적으로 주거 및 생계를 같이하는 사람의 집단으로 규정되고 있다. 최근에는 여러 부문에서 세대라는 용어를 가구나 집으로 순화하여 사용하고 있다. 이러한 측면에서만 보자면 총조사에서 사용하는 가구와 동일하고 매칭하는 데 특별한 문제가 없다.

그러나 세대라는 개념 자체가 법률적인 성격이 강하기 때문에 실제로 어떤 식으로 사용되는지를 알아볼 필요가 있다. 세대를 법률에서 규정하는 대표적인 부서가 재경부와 행정자치부이다. 먼저 재경부에서 사용하는 세대는 ‘가족’ 개념이 강한 것이다. 예를 들어, 소득세법시행령에 따르면 거주자와 배우자가 거주지를 달리하더라도 같은 세대로 본다는 것이다. 거주지와는 상관없이 가족은 실질적인 의미에서 같은 경제적 범위 안에 들어 있다는 점을 고려한 구분이다.

한편 행정자치부의 세대 개념은 주로 가족관계를 중심으로 세대를 구성한다는 점에서 재경부와 유사하다. 그러나 동일한 거주지에서 생활을 좀 더 강조하고 있다.²⁸⁾ 즉 세대를 실질적인 생활단위로 보고 있는 행정자치부는 세대 개념을 정의할 때 같이 생활한다는 측면을 강조하고 있다. 따라서 가족관계의 여부와 상관없이 거주지 중심의 세대 개념을 사용하고 있다. 그렇지만 행정자치부에서는 이러한 세대 개념을 법으로 명확하게 정의하고 있지

28) 물론 주민등록에서 하나의 세대를 형성할 때, 법률적인 가족 관계만을 허용하는 것은 아니다. ‘동거인’으로 설정할 수도 있다. 예를 들면, 고아원에 있는 아이들은 일반적으로 그리고 주민등록상으로 고아원주소로 주민등록도 설정하게 된다. 물론 간혹 어떤 경우에는 고아원 원장의 동거인으로 등재되는 경우도 있다.

않다. 주민등록법과 관련 시행령에서는 세대, 세대주, 세대 명부 등 관련 용어를 사용하고 있지만, 어떤 조항에도 이에 관한 정확한 정의를 내리고 있지 않다.

반면에 통계청에서 사용하고 있는 가구 개념은 생활 단위를 강조하고 있다. 즉 가족이나 혈연 여부와 관계없이 취사와 취침 같은 생계를 같이 할 경우 이를 한 생활단위로 보고 있다. 따라서 법률적인 호주 또는 세대주와는 관계없이 그 가구를 실질적으로 대표하는 사람이 가구주다. 만약 혈연관계가 없는 사람이 2인 이상 함께 사는 경우에는 그 중 연장자나 대표자를 선정하여 가구주로 하는 등 행정자치부의 세대주와는 선정하는 방식이 다르다. 더욱이 이런 경우 총조사에서는 1가구인 가구가 주민등록의 세대명부에는 2세대로 기록될 가능성이 매우 높다. 이러한 면에서 주민등록의 세대 개념과 총조사의 가구 개념은 동일한 거주지와 생활 단위라는 측면을 강조한다는 점에서 유사하다. 다만 주민등록의 세대는 혈연을 바탕으로 하고, 추가적으로 혈연관계가 없는 동거인 개념을 추가한 것이다. 하지만, 총조사의 가구는 좀 더 실질적인 생활단위라는 측면을 강조한다는 점에서 차이가 있을 수밖에 없다.

2. 비교 결과

가. 자료

먼저 총조사 자료로 통계청이 집계하여 홈페이지(<http://www.nso.go.kr>)에 공표하는 읍면동 수준의 인구집계자료를 이용하였다. 총조사는 2005년 11월 1일부터 15일까지 진행되었으며, 11월 1일 0시를 기준으로 인구수를 파악하여 읍면동별로 공표하고 있다. 한편 주민등록자료로는 행정자치부가 2005년 12월 31일 24시를 기준으로 주민등록에 등재된 인구수를 읍면동별로 집계하여 공표한 주민등록자료를 이용하였다. 이 두 자료는 기준 시점 자체에서 약 두 달 정도 차이가 난다.

<표 5-5> 세대와 가구 개념 비교

	세대 - 재경부	세대 - 행정자치부	가구 - 통계청
요점	‘가족’ 개념이 강함	‘생활’을 근거, 가구에 가까움	생활 단위를 뜻함
근거	<p>- 소득세법시행령 제154조 ① 제89조제1항제3호에서 "대통령령이 정하는 1세대 1주택"이라 함은 <u>거주자 및 그 배우자가 그들과 동일한 주소 또는 거소에서 생계를 같이 하는 가족과 함께</u> 구성하는 1세대(이하 "1세대"라 한다) ② 각호의 1에 해당하는 경우에는 배우자가 없는 때에도 이를 제1항의 규정에 의한 1세대로 본다. [개정 2005.12.31] 1. 당해 거주자의 연령이 30세 이상인 경우 2. 배우자가 사망하거나 이혼한 경우 3. 법 제4조의 규정에 따른 소득이 「국민기초생활보장법」 제2조제6호의 규정에 따른 최저생계비 수준 이상으로서 소유하고 있는 주택 또는 토지를 관리·유지하면서 독립된 생계를 유지할 수 있는 경우. 다만, 미성년자의 경우를 제외하되, 미성년자의 결혼, 가족의 사망 그 밖에 재정경제부령이 정하는 사유로 1세대의 구성이 불가피한 경우에는 그러하지 아니하다.</p>	<p>- 주민등록법 시행규칙 제6조 (거주지이동에 따른 관련자료의 정리등) ① 전입신고를 받은 시장·군수 또는 구청장은 영 제20조 및 영 제21조의 규정에 의하여 이송받은 주민등록표를 지체없이 정리하여야 한다. ② 시장·군수 또는 구청장이 제1항의 규정에 의한 정리를 한 때에는 영 제6조제3항의 규정에 의한 세대명부·주민등록전출자명부 및 주민등록전입자명부(이하 "세대명부등"이라 한다)를 작성하여야 한다.</p>	<p>- 가구(Household): 1인 또는 2인 이상이 모여서 취사, 취침 등 생계를 같이 하는 생활단위를 말한다. - 가구주(Household head): 호주 또는 세대주와는 관계없이 그 가구를 실질적으로 대표하는 사람을 말하며, 혈연관계가 없는 사람이 2인 이상 함께 사는 경우에는 그 중 연장자 또는 대표자를 선정하여 가구주로 한다. (인구주택총조사) 호주 또는 세대주와는 관계없이 해당가구의 구성원으로서 그 가구의 생계를 책임지고 있으며, 또한 사실상 생계유지를 위한 비용을 조달하는 사람을 가구주로 한다.(가계조사)</p>

통계청의 가구를 집계하는 기준에는 일반가구와 총가구의 두 가지가 있다. 일반가구는 총가구 수에서 집단시설, 외국인 가구 등을 제외한 가구를

의미한다. 따라서 이 절에서는 일반가구와 총가구 두 가지로 집계한 값을 제시한다.

나. 전국 수준

2005년 12월 31일 현재 주민등록의 세대수는 17,857,511이고, 2005년 11월 1일 현재 총조사의 가구수(모든 형태의 가구를 포함하는 총가구)는 15,988,274가구이다. 주민등록의 세대수가 인구주택총조사의 가구보다 약 2백만개 많은 것으로 나타났다. 이를 주민등록인구 세대 대비 총조사가구의 비율로 보면 85.9%로, 주민등록인구와 총조사인구의 차이보다 많은 것이다. 참고로, 주민등록 대비 인구주택총조사 인구의 비율은 96.4%다(이건, 2007).

이러한 차이가 나타난 이유는 여러 가지가 있을 것이다. 무엇보다도 주민등록과 총조사에서 인구수나 세대수를 파악하는 방법이 다르다. 주민등록에는 유학이나 취업 등 해외 단기 이주자가 포함되어 있다. 아울러 여러 가지 행정적, 경제적, 법률적 이유로 인해 주민등록상 분가를 하는 경우가 많다. 예를 들며, 주택 청약을 위해 단독 가구로 주민등록을 이전하는 경우도 많다. 아울러 많은 정책적, 행정적 판단을 할 때, 근거로 삼고 있는 것이 주민수인데 이럴 경우 주민등록이 되어 있는 주민 수를 사용하는 경우가 많다.²⁹⁾

주민등록과 총조사의 이러한 차이를 좀 더 체계적으로 살펴보기 위해서 앞 절에서 같은 개념의 차이율을 이용한다. 앞 절과의 차이는 인구 대신 가구와 세대를 사용한다는 점이다. 차이율은 주민등록세대를 기준으로 하고 주민등록세대수와 총조사가구수의 차이를 계산한 것이다. 식으로 표시하면 “차이율=[(주민등록세대수 - 총조사가구수) / 주민등록세대수] * 100”이다. 이렇게 계산한 전국 수준의 차이율은 총가구가 10.5%이고 일반가구는 11.0%이다(<표 5-6>).

29) 따라서 여러 가지 이유에 의한 차이를 좀 더 신속하게 수정하고 체계화하기 위해서 앞으로 제도적인 측면을 보완할 필요가 있다.

<표 5-6> 주민등록세대와 총조사가구의 시도별 차이

시도	주민등록 세대	총조사 총가구	총조사 일반가구	차이율(%) 총가구	차이율(%) 일반가구
전국	17,857,511	15,988,274	15,887,128	10.5	11.0
서울특별시	3,871,024	3,341,352	3,309,890	13.7	14.5
부산광역시	1,270,612	1,190,107	1,186,378	6.3	6.6
대구광역시	865,766	817,620	814,585	5.6	5.9
인천광역시	933,686	828,012	823,023	11.3	11.9
광주광역시	481,935	461,387	460,090	4.3	4.5
대전광역시	505,650	480,466	478,865	5.0	5.3
울산광역시	365,197	340,652	339,095	6.7	7.1
경기도	3,910,886	3,361,657	3,329,177	14.0	14.9
강원도	574,019	522,225	520,628	9.0	9.3
충청북도	547,213	507,541	505,203	7.2	7.7
충청남도	756,073	663,445	659,871	12.3	12.7
전라북도	683,683	621,735	619,958	9.1	9.3
전라남도	751,732	667,962	666,319	11.1	11.4
경상북도	1,006,080	943,135	938,840	6.3	6.7
경상남도	1,129,320	1,061,259	1,056,007	6.0	6.5
제주도	204,635	179,719	179,199	12.2	12.4

* 차이율 = {(주민등록세대-총조사가구)/주민등록세대} * 100

다. 시도 수준

<표 5-6>에는 16개 시도별로 비교한 주민등록세대수와 총조사가구수를 제시하였다. 전국 수준과 마찬가지로 차이율을 기준으로 시도별 차이를 살펴 보면, 지역별 차이가 드러나고 있다. 전국을 기준으로 볼 때, 총가구의 경우 차이율이 10.5%이지만, 시도 수준에서 주민등록세대수와 총조사의 가구 대한

차이율의 평균은 8.8%이, 일반가구에 대한 차이율의 평균은 9.2%이다.

전반적으로 시도 수준에서 주민등록세대수와 총조사가구수 차이 유형은 개별 인구의 경우와 유사하다. 서울특별시, 인천광역시, 경기도와 같은 수도권과 충청남도, 제주도의 경우는 평균보다 다소 높다. 일반가구의 경우도 이와 유사하다. 이들 지역은 다른 지역에 비해 지역 개발이 상대적으로 활발하게 이루어지고 있다는 공통점이 있다. 수도권의 경우 끊임없이 인구가 유입되고 있고, 충청남도는 행복도시가 제주도의 경우 자치도로 인해 최근에 지역개발이 활기를 띠고 있는 지역이다.

반면에 평균보다 차이율이 낮은 지역은 인천을 제외한 나머지 광역시 지역이다. 여기에는 대구광역시, 광주광역시, 대전광역시, 울산광역시가 포함되어 있다. 광역시를 제외한 시도 지역 중에서 차이율이 상대적으로 낮은 지역은 경상도 지역이다. 경상북도와 경상남도의 경우 총가구 차이율이 각각 6.3%와 6.0%로 나타난다. 부산광역시와 울산광역시를 포함하면 전체 영남 지역에서 상대적으로 낮은 차이율을 보인다.

라. 시군구 수준

(그림 5-7)과 (그림 5-8)은 시군구 수준에서 주민등록세대수와 총조사가구수를 총가구와 일반가구로 나누어 차이율의 분포를 각각 보여주고 있다. 먼저 전반적인 평균값을 살펴보면, 시군구 수준에서는 주민등록세대수와 총조사가구수 차이율은 시도 수준에 비해 높다. 총가구의 경우 평균은 11.5%이고, 일반가구의 경우 평균 차이율은 12.0%로 나타난다.

그 다음으로 그러한 차이율이 어떠한 분포를 보이는가를 살펴보자. 표준편차로 보면, 총가구의 경우 표준편차는 5.4이고, 일반가구의 경우 평균 차이율은 5.5로 나타난다. 이렇게 평균값과 표준편차를 비교해 볼 때, 시도 수준에서 나타나는 특징이 시군구 수준에서도 유사하게 나타난다. 다만 시도 수준보다는 시군구 수준에서 분포가 넓게 퍼져있다.

이러한 시군구별 차이율을 구체적으로 살펴보자. <표 5-7>과 <표 5-8>은 총가구와 일반가구의 시군구별 차이율을 총조사가구가 많은 경우와 총조

사가구가 적은 경우로 나누어 보여주고 있다.

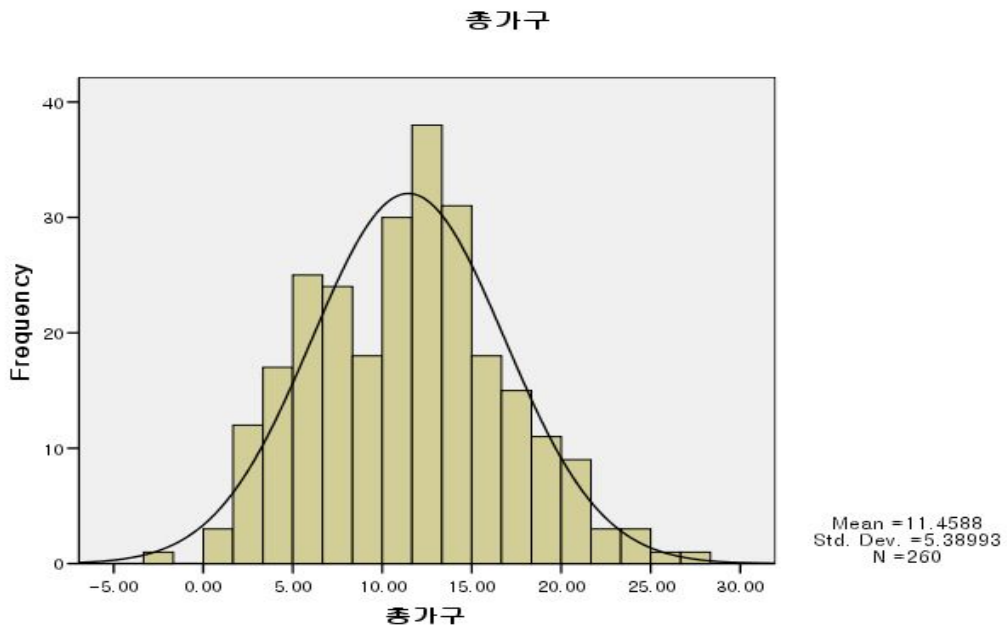
먼저 총가구의 경우를 살펴보자. 총조사가구가 많은 곳은 경상시가 유일하다. 차이율이 -1.9%로 나타났다.³⁰⁾ 경상북도 경산시는 비교적 최근에 신도시 개발이 된 곳이다. 이곳에는 영남대학교, 대구대학교, 대구가톨릭대학교, 대구한의대학교, 경일대학교 등 각종 대학이 위치해 있다.

반면에 총조사가구가 적은 곳 중에서 차이율이 20%가 넘는 곳만을 살펴보자. 부산광역시 강서구의 경우 차이율이 27.1%에 이른다. 그 다음으로 높은 지역은 인천광역시 옹진군으로, 차이율이 25.5%이다. 다음으로 경기도 하남시, 파주시 등 수도권 지역과 강원도 화천군과 경기도 연천군, 전라남도 장성군 등 군부대가 많이 있는 지역의 차이율이 대체로 높은 편이다. 제주도 북제주군도 차이율이 22.2%에 이른다.³¹⁾

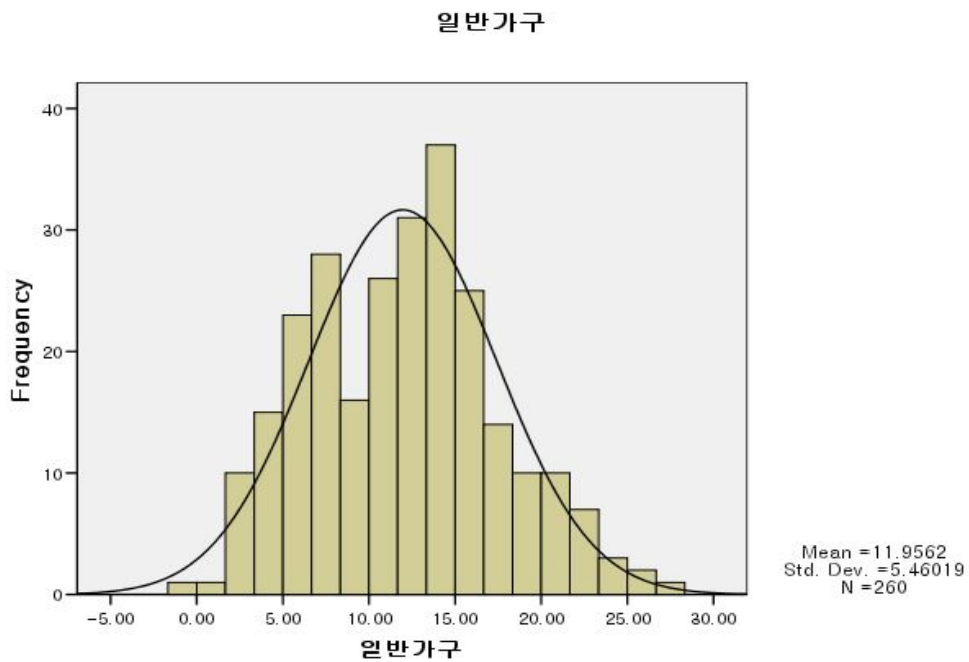
일반가구의 경우에서도 전체적인 흐름은 유사하게 나타난다. 총조사가구가 많은 곳은 총가구의 경우와 마찬가지로 경상북도 경산시가 유일하다. 총조사가구가 적은 곳 중에서 차이율이 20%가 넘는 지역을 살펴보면, 차이율이 가장 높은 지역은 28.1%로 부산광역시 강서구이다. 그 다음으로는 인천광역시 옹진군으로 차이율이 25.7%에 이른다. 총가구의 경우와 일반가구의 경우에서 차이가 나는 서울특별시 종로구이다. 총가구에서는 상위 10위 안에 포함되지 않았던 서울특별시 종로구가 포함되었다. 종로구의 차이율은 22.1%에 이른다.

30) 차이율이 상대적으로 낮은 지역은 시도별 차이에서 알 수 있듯이 각 광역시와 영남지역의 시군구 지역이다. 차이율이 가장 낮은 지역은 총조사가구수가 주민등록세대수보다 많은 지역인 경상북도 경산시이다. 그 다음으로 낮은 지역은 대전광역시 서구, 경상남도 거제시, 경상북도 안동시로, 차이율이 각각 1.0%, 1.2%이다. 이들 지역에서는 적어도 세대수와 가구수의 차이가 거의 없는 것으로 나타났다.

31) 주민등록법시행령에서는 군인의 경우 다음과 같이 주민등록에 관한 규칙을 제정해 놓고 있다. “제1항의 등록에 있어서 영내에 기거하는 군인에 대하여는 그가 속하는 세대의 거주지에서 본인 또는 그 세대주의 신고에 의하여 등록하여야 한다.”



(그림 5-7) 주민등록세대와 총조사가구의 시군구별 차이율-총가구



(그림 5-8) 주민등록세대와 총조사가구의 시군구별 차이율-일반가구

<표 5-7> 총조사가구가 주민등록세대보다 20% 이상 적은 경우-총가구

순번	시/도	시/군/구	총조사 총가구수	주민등록 세대수	차이율 (%)
1	부산광역시	강서구	21,311	15,530	27.1
2	인천광역시	옹진군	6,801	5,068	25.5
3	경기도	하남시	50,859	38,492	24.3
4	경기도	파주시	100,278	76,388	23.8
5	강원도	화천군	9,731	7,433	23.6
6	경기도	양평군	34,641	26,613	23.2
7	전라남도	장성군	20,228	15,642	22.7
8	제주도	북제주군	36,886	28,709	22.2
9	경기도	가평군	21,443	16,830	21.5
10	경기도	연천군	18,377	14,494	21.1
11	전라북도	진안군	11,622	9,189	20.9
12	전라남도	담양군	19,976	15,872	20.5
13	서울특별시	종로구	69,982	55,748	20.3
14	제주도	남제주군	27,083	21,578	20.3
15	경기도	포천시	59,793	47,688	20.2
16	인천광역시	중구	38,707	30,889	20.2

* 차이율 = {(주민등록세대-총조사가구)/주민등록세대} * 100

<표 5-8> 총조사가구가 주민등록세대보다 20% 이상 적은 경우-일반가구

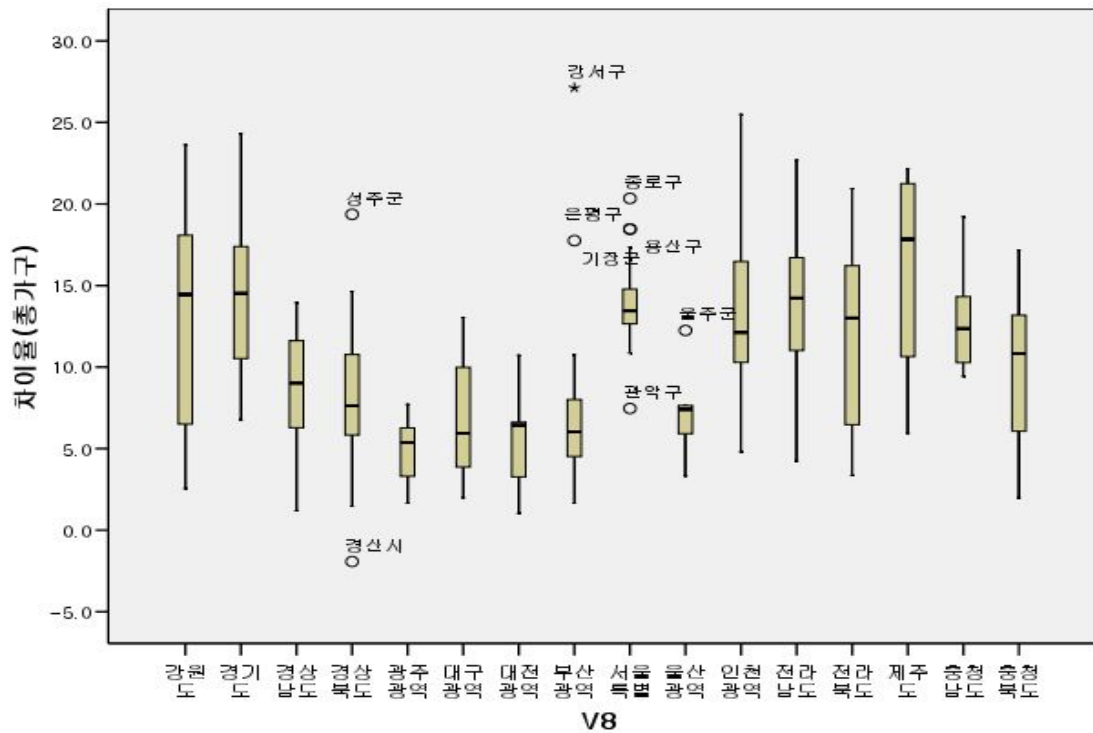
순번	시/도	시/군/구	총조사 일반가구수	주민등록 세대수	차이율 (%)
1	부산광역시	강서구	21,311	15,315	28.1
2	인천광역시	옹진군	6,801	5,052	25.7
3	경기도	하남시	50,859	38,160	25.0
4	경기도	파주시	100,278	75,654	24.6
5	강원도	화천군	9,731	7,416	23.8
6	경기도	양평군	34,641	26,496	23.5
7	전라남도	장성군	20,228	15,531	23.2
8	제주도	북제주군	36,886	28,600	22.5
9	서울특별시	종로구	69,982	54,534	22.1
10	경기도	가평군	21,443	16,754	21.9
11	경기도	연천군	14,494	18,377	21.8
12	경기도	포천시	47,688	59,793	21.8
13	서울특별시	용산구	79,340	97,324	21.8
14	경기도	광주시	63,376	79,288	21.1
15	인천광역시	중구	30,889	38,707	21.1
16	전라북도	진안군	9,189	11,622	21.0
17	경기도	양주시	46,291	58,570	21.0
18	경기도	김포시	59,249	74,942	20.9
19	전라남도	담양군	15,818	19,976	20.8
20	제주도	남제주군	21,528	27,083	20.5
21	경기도	분당구	130,722	163,608	20.1
22	경상북도	성주군	15,197	18,999	20.0
23	충청남도	아산시	68,263	85,308	20.0

* 차이율 = {(주민등록세대-총조사가구)/주민등록세대} * 100

이러한 시군구별 차이를 분포와 함께 고려해 보면, 주민등록세대수와 총조사가구수 차이를 다른 측면에서 살펴볼 수 있다. 여러 가지 방법 중에서, 여기에서는 상자도표(box plot)를 사용하고자 한다. 일반적으로 상자도표를 통해 우리는 분포와 평균을 시각적으로 나타내서 이상값을 찾아낼 수 있다.

이를 위해 시군구별 주민등록세대수와 총조사의 가구수의 차이를 상자도표를 작성하였다((그림 5-9)).

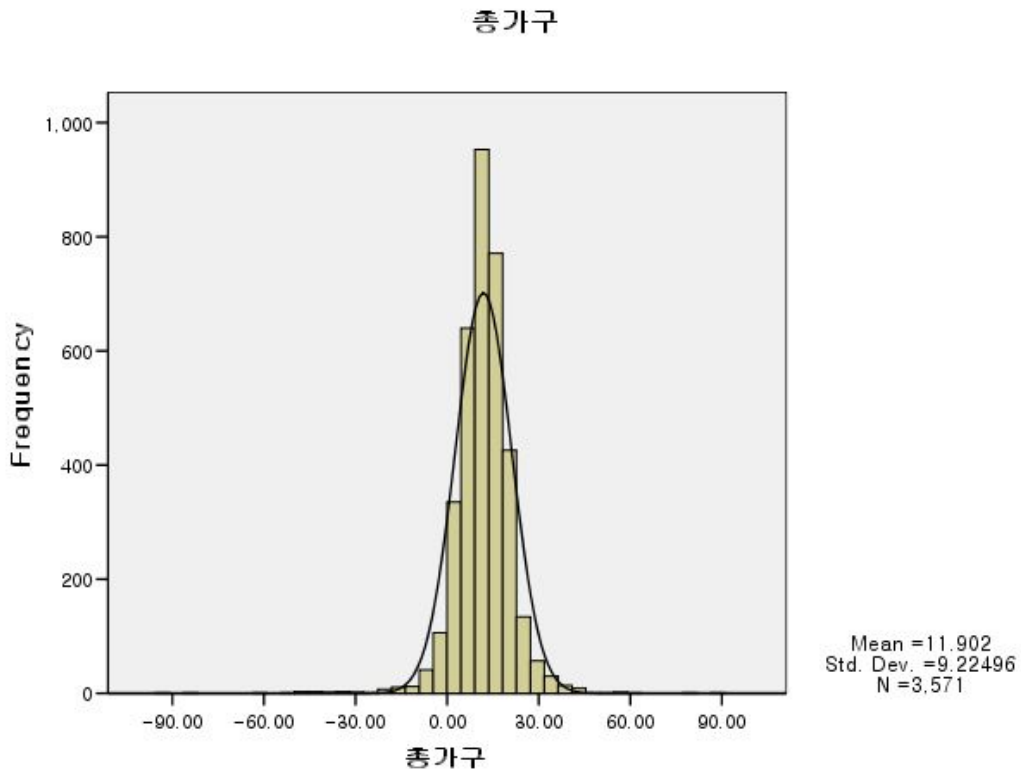
상자도표에서 나타난 개별적인 특이 사례를 살펴보면, 시도별로 볼 때 이상값이 나타난 곳은 경상북도, 부산광역시, 울산광역시, 서울특별시이다. 경상북도의 경우 성주군과 경산시가 부산광역시의 경우 기장군과 강서구가 이상값으로 나타났다. 성주군은 차이율이 아주 높게 나타났다면 경산시는 차이율이 음수로 나타났다. 부산광역시의 경우 강서구와 기장군이 평균보다 훨씬 높은 차이율을 보이고 있다. 울산광역시의 경우 울주군의 차이율이 평균보다 높게 나타난다. 마지막으로 서울시의 경우는 종로구, 용산구, 은평구, 관악구가 특이한 사례로, 모두 평균보다 높은 차이율을 보이고 있다.



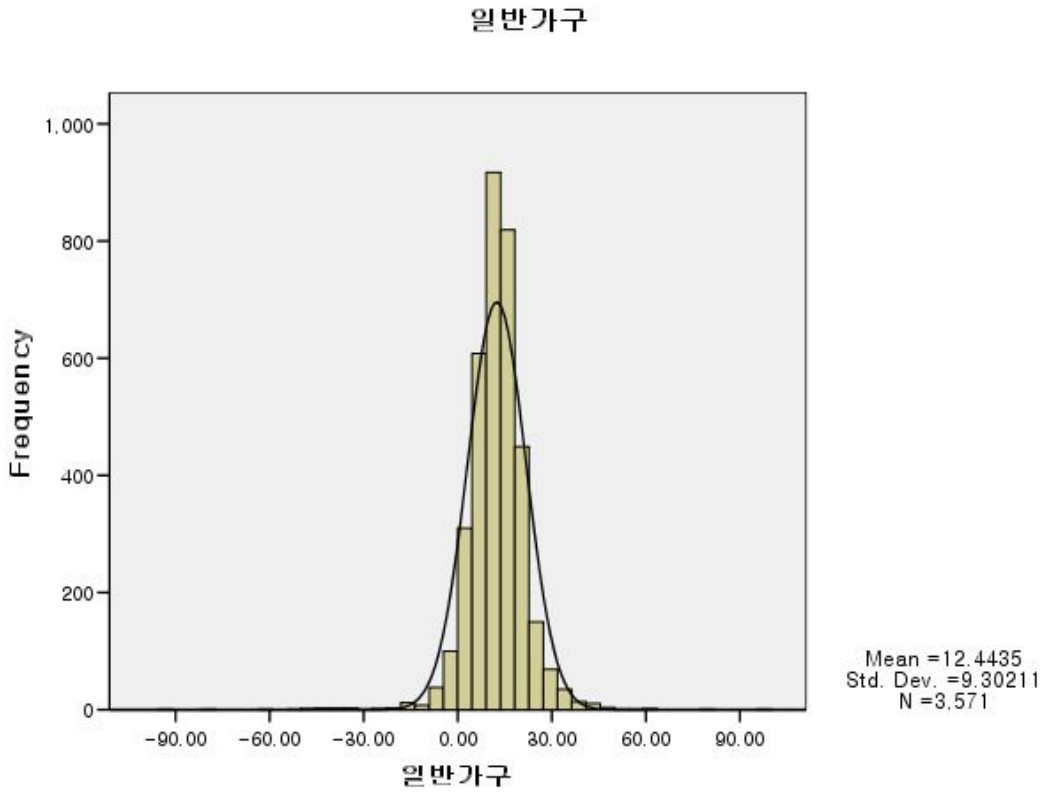
(그림 5-9) 주민등록세대와 총조사가구의 시군구별 차이율 상자도표-총가구

마. 읍면동 수준

마지막으로 읍면동 수준에서 주민등록세대수와 총조사가구수를 비교해 보자. 읍면동 같은 소지역에서는 차이가 상대적으로 크게 나타난다. (그림 5-10)과 (그림 5-11)는 읍면동 수준에서 총가구와 일반가구의 차이율 분포를 제시하고 있다. 전체적으로 총가구와 일반가구 유사한 결과를 보여주고 있다. 일반가구의 평균 차이율과 표준편차가 총가구의 평균 차이율과 표준편차에 비해 다소 높다. 구체적으로 평균 차이율을 살펴보면, 총가구의 경우 평균 차이율은 11.9%이고, 일반가구의 경우 평균 차이율은 12.4%로 나타난다. 그리고 차이율 분포의 경우는 총가구의 경우 표준편차는 9.2이고, 일반가구의 경우 차이율의 표준편차 9.3으로 나타난다.



(그림 5-10) 주민등록세대와 총조사가구의 읍면동별 차이율-총가구



(그림 5-11) 주민등록세대와 총조사가구의 읍면동별 차이율-일반가구

여기에서 주민등록세대와 총조사가구 사이에 나타나는 차이의 분포를 살펴볼 필요가 있다. <표 5-9>는 간단하게 주민등록세대 대비 총조사가구 비율 분포를 제시한 표이다. 주민등록세대 대비 총조사가구 비율이 85%미만인 경우는 총가구의 경우는 34.6%이고, 일반가구의 경우는 37.3%로 나타났다. 주민등록세대 대비 총조사가구 비율이 85%이상 90%미만인 경우는 총가구가 28.9%, 일반가구가 28.6%로 나타났다. 한편 주민등록세대와 총조사가구 수의 차이가 $\pm 5\%$ 이내로 나타나는 (비율이 95%부터 105% 사이에 있는) 읍면동 지역은 전체 중에서 총가구의 경우는 13.5%, 일반가구의 경우는 12.8%로 나타났다.

<표 5-9> 읍면동 수준의 총조사가구/주민등록세대 분포

총조사가구/주민등록세대비율(%)	총가구		일반가구	
	빈도	%	빈도	%
85 미만	1,236	34.6	1,332	37.3
85-90	1,033	28.9	1,022	28.6
90-95	740	20.7	690	19.3
95-100	370	10.4	352	9.9
100-105	111	3.1	102	2.9
105 초과	80	2.2	72	2.0
전체	3,570	100	3,570	100

그러면 구체적으로 어느 지역에서 차이율이 크거나 작은지를 살펴보기로 하자. <표 5-10>과 <표 5-11>은 총가구의 읍면동별 차이율과 일반가구의 읍면동별 차이를 각각 보여주고 있다. 먼저 총조사가구가 주민등록세대보다 많은 경우를 살펴보자. 이 경우에 해당하는 지역으로는 충청북도 충주시 달천동, 경상북도 경산시 북부동, 경기도 안성시 대덕면, 강원도 삼척시 정라동 등이 있다.

반면에 총조사가구가 주민등록세대보다 적은 지역으로는 서울특별시 송파구 잠실1동과 성북구 월곡3동, 충청북도 청주시 흥덕구 강서2동, 경기도 파주시 월롱면, 충청남도 아산시 탕정면 등이 있다. 이들 지역에서는 차이율이 50.0%를 넘는다. 특히 잠실1동의 경우는 차이율이 90.9%로 나타난다.

이러한 차이율 분포의 원인을 검토할 필요가 있다. 먼저 총조사가구가 많은 경우는 대학 주변 지역이라는 특징이 있다. 달천동의 경우 주변에 건국대학교 충주캠퍼스가, 북부동에는 영남대학교가, 대덕면에는 중앙대학교 안성캠퍼스, 신림9동에는 서울대학교가 위치해 있다. 반면에 총조사가구가 적은 경우는 재개발이거나 기업도시 등 지역개발이 활발하게 이루어지고 있다.

<표 5-10> 주민등록세대와 총조사가구의 읍면동별 차이율-총가구

	순위	시도	시/군/구	읍면동	주민등록 세대	총조사 총가구	차이율(%) 총가구
총조사 가구가 많은 경우	1	충청북도	충주시	달천동	2,155	4,148	(-92.5)
	2	경상북도	경산시	북부동	5,739	10,445	(-82.0)
	3	경기도	안성시	대덕면	4,104	6,624	(-61.4)
	4	강원도	삼척시	정라동	2,965	4,495	(-51.6)
	5	강원도	강릉시	경포동	2,229	3,328	(-49.3)
	6	강원도	원주시	홍업면	2,771	4,102	(-48.0)
	7	전라북도	익산시	신동	6,871	10,120	(-47.3)
	8	인천광역시	남구	용현4동	3,435	4,971	(-44.7)
	9	서울특별시	관악구	신림9동	12,053	17,335	(-43.8)
	10	강원도	강릉시	내곡동	3,893	5,510	(-41.5)
총조사 가구가 적은 경우	1	서울특별시	송파구	잠실1동	11	1	90.9
	2	서울특별시	성북구	월곡3동	346	78	77.5
	3	충청북도	청주시 흥 덕구	강서2동	2,399	911	62.0
	4	경기도	파주시	월롱면	4,576	1,912	58.2
	5	충청남도	아산시	탕정면	8,304	3,750	54.8
	6	경기도	시흥시	과림동	1,191	587	50.7
	7	인천광역시	중구	용유동	1,729	903	47.8
	8	대전광역시	동구	중앙동	1,850	1,015	45.1
	9	경기도	성남시 분 당구	관교동	370	204	44.9
	10	전라북도	군산시	옥도면	1,578	879	44.3

* 차이율 = {(주민등록세대-총조사가구)/주민등록세대} * 100

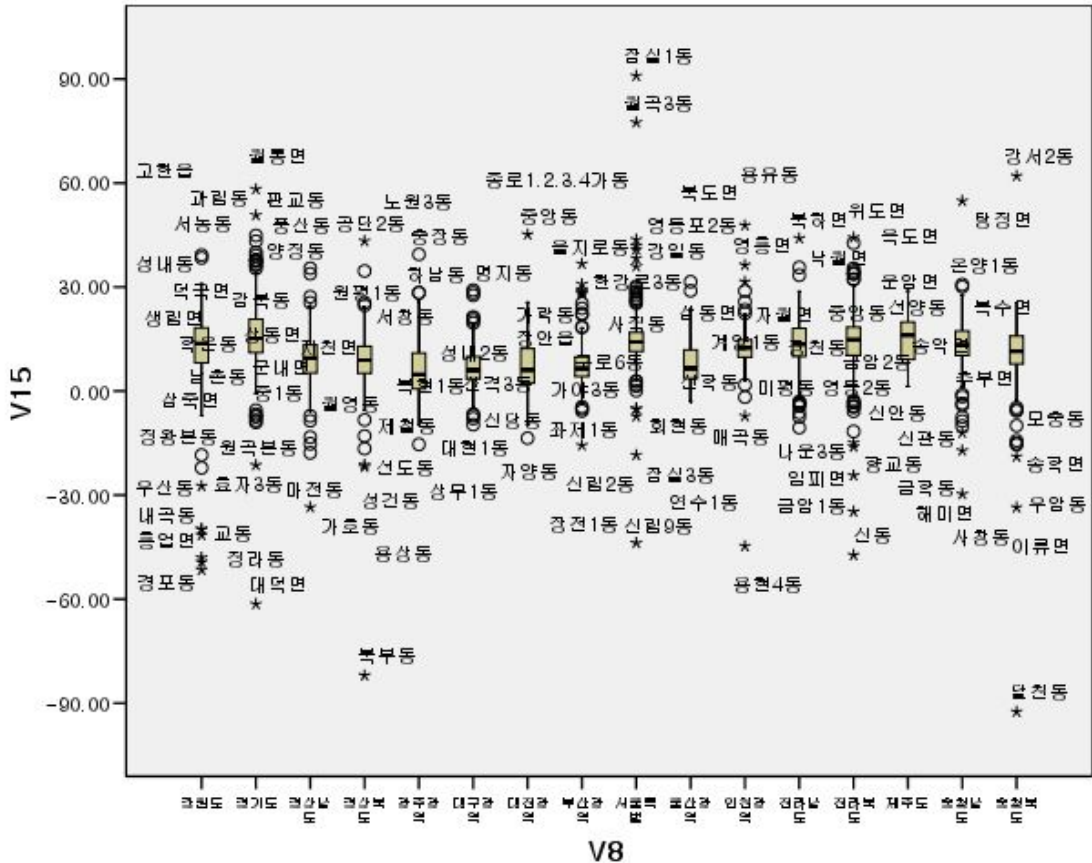
<표 5-11> 주민등록세대와 총조사가구의 읍면동별 차이율-일반가구

	순위	시도	시/군/구	읍면동	주민등록세대	총조사일반가구	차이율(%) 총가구
총조사 가구가 많은 경우	1	충청북도	충주시	달천동	2,155	4,127	(-91.5)
	2	경상북도	경산시	북부동	5,739	10,343	(-80.2)
	3	경기도	안성시	대덕면	4,104	6,542	(-59.4)
	4	강원도	삼척시	정라동	2,965	4,491	(-51.5)
	5	강원도	강릉시	경포동	2,229	3,237	(-45.2)
	6	강원도	원주시	홍업면	2,771	4,056	(-46.4)
	7	전라북도	익산시	신동	6,871	10,030	(-46.0)
	8	인천광역시	남구	용현4동	3,435	4,906	(-42.8)
	9	서울특별시	관악구	신림9동	12,053	17,267	(-43.3)
	10	강원도	강릉시	내곡동	3,893	5,479	(-40.7)
총조사 가구가 적은 경우	1	서울특별시	송파구	잠실1동	11	0	100.0
	2	서울특별시	성북구	월곡3동	346	78	77.5
	3	충청북도	청주시 흥덕구	강서2동	2,399	899	62.5
	4	경기도	파주시	월롱면	4,576	1,848	59.6
	5	충청남도	아산시	탕정면	8,304	3,697	55.5
	6	경기도	시흥시	과림동	1,191	558	53.1
	7	인천광역시	중구	용유동	1,729	902	47.8
	8	대전광역시	동구	중앙동	1,850	998	46.1
	9	경기도	성남시 분당구	관교동	370	197	46.8
	10	전라북도	군산시	옥도면	1,578	869	44.9

* 차이율 = {(주민등록세대-총조사가구)/주민등록세대} * 100

개별 읍면동의 실태도 주민등록세대원수와 총조사 인구수 차이의 경우와 유사하다. 마찬가지로 상자도표를 통해 이상값을 살펴보기로 하자. (그림 5-12)은 읍면동 수준에서 이상값을 보여주고 있다. 읍면동과 같은 소지역에서는 재개발 등 각종 지역사회 개발에 큰 영향을 받고 있다. 예를 들면, 서울시의 경우 잠실1동, 잠실3동, 월곡3동, 신림9동 등이 차이가 많이 나는 지

역이다.경기도의 경우는 관교동, 대덕면 등이고, 충청남도의 경우 탕정면, 충청북도의 경우 달천동이 차이가 많이 나는 지역이다. 서울과 판교의 경우는 재개발과 관련이 깊고, 탕정면은 기업도시, 달천동은 대학 근처 지역이라는 특징 있다.



(그림 5-12) 주민등록세대와 총조사가구의 읍면동별 차이율 상자도표-총가구

3. 소결

이 장에서는 주민등록세대수와 총조사가구수를 비교할 경우 어떤 문제가 있는지를 알아보았다. 이를 위해 세대와 가구의 개념을 알아보고, 세대수와 가구수를 시도, 시군구, 읍면동 세 단계로 나누어 살펴보았다.

먼저 주민등록자료의 세대와 총조사의 가구의 개념적인 정의와 실제자료

수집을 고려할 때, 충분히 비교가 가능한 개념이다. 두 개념 모두 생활단위라는 측면에서 유사하다. 물론 주민등록과 총조사 자체의 이론적인 목적과 실질적인 용도가 동일하지 않기 때문에 두 개념 사이에 다소 차이는 발생할 수 있다. 그러나 다른 정부 부처, 예를 들면 재경부에서 사용하는 세대 개념 보다는 훨씬 비교 가능성이 높다.

이러한 개념적인 검토를 기반으로, 각 수준별로 주민등록 자료와 총조사 자료를 검토하였다. 먼저 전국 수준에서 차이율은 총가구가 10.5%이고 일반가구는 11.0%이다. 이는 주민등록세대원수와 총조사 인구수의 차이보다 높은 수치이다. 이러한 차이율은 시도, 시군구, 읍면동 수준으로 내려갈수록 커지고 있다. 시도 수준에서는 총가구의 경우 차이율이 가장 낮은 지역의 차이율이 4.3%인 반면에 가장 높은 지역의 차이율은 14.0%로 나타났다. 시군구 수준에서는 총가구의 경우 차이율이 가장 높은 지역의 차이율이 27.1%에 이르렀다. 마지막으로 읍면동 수준에서 총가구의 차이율이 가장 높은 지역의 차이율이 90.9%로 나타났다.

이렇게 현재 상태에서는 주민등록세대수로 총조사의 가구를 예측하기는 쉽지 않다. 앞으로 주민등록 같은 행정자료도 인구통계를 집계하기 위해서는 몇 가지 종류의 연구가 필요하다. 무엇보다도 두 종류의 자료 비교를 위한 상당히 많은 연구와 시계열 자료에 대한 분석이 필요하다. 현재 연구에서는 최근 자료에 대한 연구가 이루어 졌다. 그러나 이러한 자료에 대한 지속적인 연구를 통해 좀 더 지속적인 경향을 파악할 수 있을 것이다. 둘째, 총조사 자료를 주민등록 자료로 대체하기 위해서는 총조사 자료뿐만 아니라 주민등록 자료 자체에 대한 연구도 진행되어야 한다. 사실 주민등록 자료의 정확성에 관한 연구가 충분하지 않다. 아울러 주민등록 자료의 수집과 이용에 있어서 발생하고 있는 현실적인 문제도 검토해야 할 필요가 있다. 따라서 주민등록 자료의 품질을 향상시키기 위한 법률적, 행정적 개선 노력도 동시에 진행될 필요가 있다.

제4절. 연령별 비교

1. 서론

이 장에서는 주민등록과 인구주택총조사의 차이를 연령별로 비교하여 분석하였다. 주민등록과 인구주택총조사의 연령별 비교가 가지는 의의들은 다음과 같다. 첫째, 주민등록과 인구주택총조사의 차이를 추정하고 주민등록의 통계적 활용가능성을 평가한다. 연령별 인구 집계는 인구통계의 핵심 항목이다. 주민등록을 통한 각 연령층 인구 집계와 인구주택총조사의 연령별 인구 집계 사이의 차이는 그 자체로서 주민등록을 통계적으로 활용할 수 있는 정도를 판단하는 중요한 지표가 된다.

또한 주민등록과 인구주택총조사 사이의 차이에 대한 연령별 및 지역수준별 비교 결과를 종합적으로 고려함으로써 주민등록의 활용 가능성을 더욱 정확히 추정할 수 있다. 전국적으로 주민등록으로 집계된 인구수는 인구주택총조사에 의해 집계된 인구수보다 약 174만 명이 많고 이 차이는 전체 인구수의 약 3.6%에 해당한다(이건, 2007). 하지만 연령대별 차이율을 고려하지 않고 전체 수준에서만 차이율을 기준으로 삼을 경우 주민등록과 인구주택총조사 사이의 차이를 과소평가할 위험이 있다. 연령대나 지역별 차이가 서로 다른 방향으로 나타나서 차이들이 상쇄될 수 있기 때문이다.

둘째, 주민등록과 인구주택총조사 사이의 연령대별 차이를 지역수준별로 파악함으로써 오차 발생 현황과 원인에 대한 분석이 가능해진다. 연령대별 차이를 지역별로 비교하고 연령대별 차이율이 높은 지역들의 특성을 파악함으로써 오차의 발생원인을 어느 정도 파악할 수 있을 것이다. 오차가 규명되면 이의 개선을 위한 조치들도 제시 가능하다.

이 연구에서 사용한 인구주택총조사와 주민등록의 연령별 집계자료는 모두 통계청이 홈페이지(<http://www.nso.go.kr>)에 공표한 읍면동 수준의 인구 집계이다. 지역수준은 전국 수준, 광역시도, 시군구, 읍면동의 4가지로 나누어서 분석을 진행하였다. 통계청이 제시한 연령별 집계는 5세 연령층을 단위로 제시되었다. 이 연구에서도 5세 연령층을 기본 단위로 연령별 비교를 수

행하였다. 모든 지역수준에서 연령층의 단위를 동일하게 사용하였으나 읍면동 수준에서는 70세 이상의 연령대를 통합하여 측정하였다. 각 지역 내 해당 연령대 인구 규모가 적기에 약간의 변화로도 차이율의 값이 크게 영향을 받기 때문이다.

2. 비교 결과

가. 전국 수준

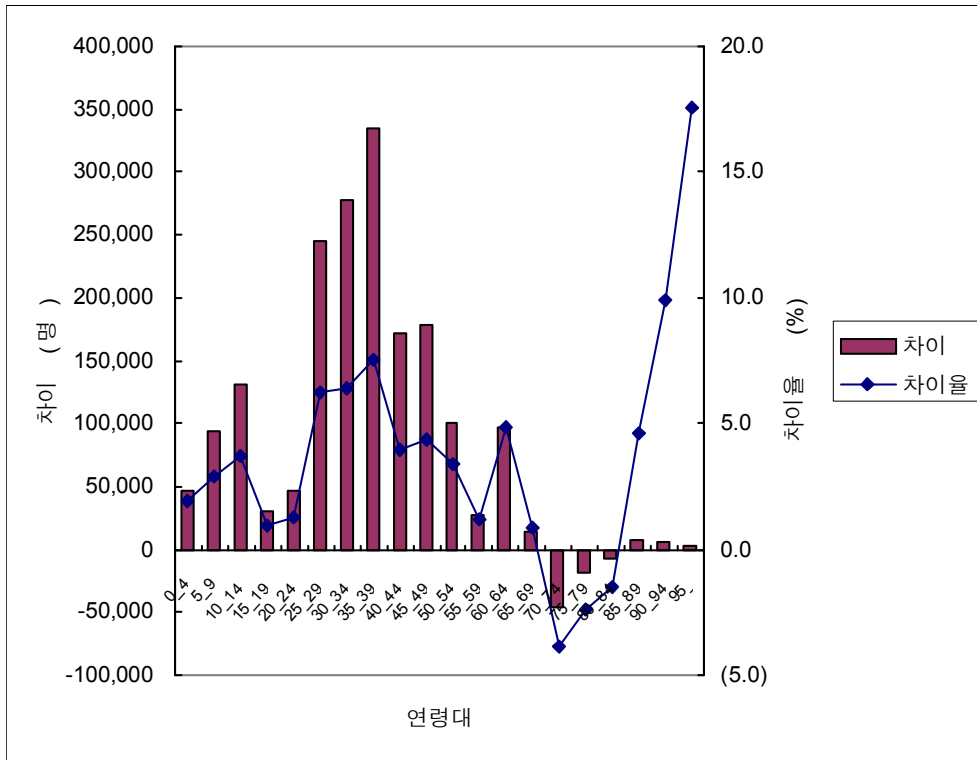
먼저 전국 수준에서 주민등록인구와 총조사인구의 차이를 연령대별로 비교하면 <표 5-12>에서 보듯이 차이와 차이율이 연령대에 따라서 크게 차이가 났다. 연령대별 차이율을 선으로 연결하면 (그림 5-13)에서 보듯이 연령대의 구간 구간별로 일정한 패턴을 가지는 것을 볼 수 있다. 차이율이 연령대에 따른 특정한 유형을 갖고 있다는 사실은 차이의 발생 원인에 대한 해석의 가능성을 제시한다.

우선 0세에서부터 14세까지는 연령의 증가에 따라 차이율이 점점 높아지는 양상이 나타난다. 다음으로 15-24세까지의 연령대에는 이전 연령대에 비해서 차이율이 감소하여 1% 내외의 차이율을 보이다가 25세 이상의 연령대에서는 차이율이 크게 높아진다. 차이와 차이율을 모두 고려할 때 35-39세 때 전체 연령구간에서 가장 큰 차이가 나타난다. 이 연령대에서는 주민등록인구가 인구주택총조사의 인구보다 약 33만 명이나 많고 차이율도 7.5%나 되어서 전체 차이율의 약 두 배 이상이나 되는 높은 수준을 보였다. 40세 이후의 연령대에서는 차이율이 낮아지는데 55-59세의 연령급간에서는 차이율이 1.2%, 65-69세의 연령급간에서는 차이율이 0.85%로 주민등록과 인구주택총조사의 차이가 상당히 적다. 이후 70-84세의 연령층에서는 부의 차이율이 나타난다. 주민등록인구에 비해서 총조사인구가 더 적은 다른 연령층과는 달리 70-84세의 노년층에서는 총조사인구에 비해 주민등록인구가 적다. 반면 85세 이상의 노년층에서는 다시 총조사인구가 주민등록인구수에 못 미치면서 정의 차이율이 높은 수준에 이른다. 90-94세에서는 차이율이 9.9%에 이르고

95세 이상에서는 차이율이 17.6%나 되어서 전체 연령층에서 가장 높은 수준이다. 하지만 이 연령대의 인구수 자체가 많지 않아서 차이의 절대 규모는 크지 않다.

<표 5-12> 연령대별 주민등록인구와 총조사인구의 차이와 차이율

연령대	주민등록	총조사	차이	차이율
0_4	2,429,599	2,382,350	47,249	1.9
5_9	3,262,179	3,168,887	93,292	2.9
10_14	3,566,171	3,434,891	131,280	3.7
15_19	3,130,806	3,100,523	30,283	1.0
20_24	3,709,318	3,662,123	47,195	1.3
25_29	3,917,226	3,671,847	245,379	6.3
30_34	4,374,750	4,096,282	278,468	6.4
35_39	4,447,026	4,112,785	334,241	7.5
40_44	4,294,353	4,123,041	171,312	4.0
45_49	4,078,848	3,900,899	177,949	4.4
50_54	2,955,991	2,855,297	100,694	3.4
55_59	2,306,049	2,278,438	27,611	1.2
60_64	1,985,434	1,888,853	96,581	4.9
65_69	1,694,460	1,680,067	14,393	0.9
70_74	1,206,036	1,252,734	(46,698)	(3.9)
75_79	748,702	766,870	(18,168)	(2.4)
80_84	425,870	432,259	(6,389)	(1.5)
85_89	179,760	171,521	8,239	4.6
90_94	56,312	50,738	5,574	9.9
95_	13,384	11,029	2,355	17.6
합 계	48,782,274	47,041,434	1,740,840	3.6



(그림 5-13) 연령대별 차이와 차이율

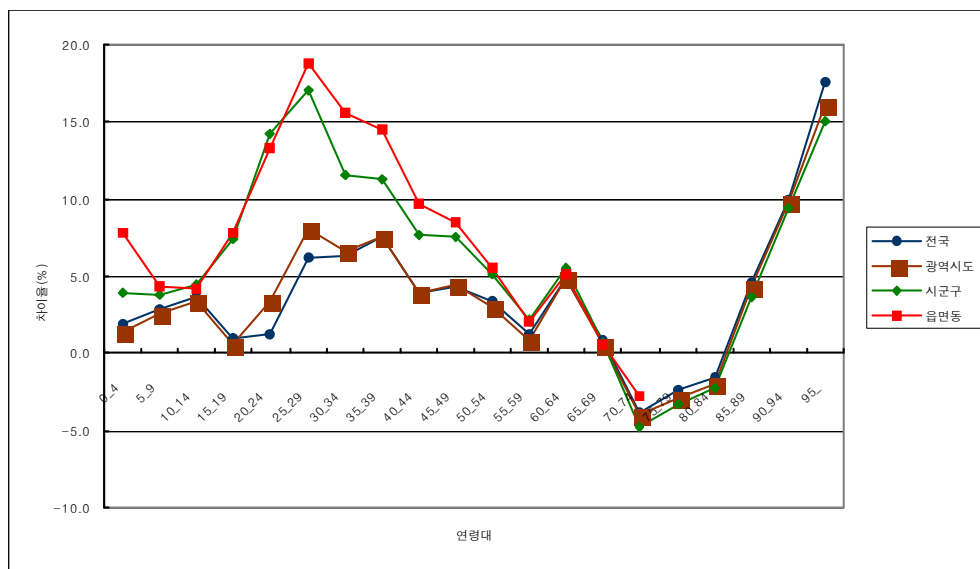
이상 연령대별 주민등록인구와 총조사의 차이를 보았을 때 다음과 같은 사실들을 발견할 수 있다. 우선 전체 인구수준에서 주민등록인구와 총조사인구를 비교할 때 차이율이 3.6%인데 연령대별로 차이율의 편차가 크게 나타나서 4% 대 이상의 높은 차이율을 보이는 연령대도 적지 않다. 또한 주민등록인구보다 총조사인구가 더 많은 연령층도 있어서 연령대들 사이에서 차이율의 상쇄효과가 발생한다.

전체적으로 생애주기에 따른 사회경제적 활동 및 인구이동의 변이와 인구주택총조사의 조사 가능성에 의해서 영향을 받은 것으로 해석된다. 예외적이라면 15-24세까지 연령기의 높지 않은 차이율이다. 학업을 수행하고 사회적 진출을 준비하는 기간의 연령층이지만 총조사를 통한 인구추정과 주민등록인구수가 비교적 근접하였다. 이후 20대 후반부터 30대까지는 경제적 사회적 활동이 가장 왕성하고 공간적 이동이 많은 시기로서 주민등록과 인구주택총조사의 차이가 크게 나타났다. 이 집단은 직장 근무지 때문에 발생하는 단기

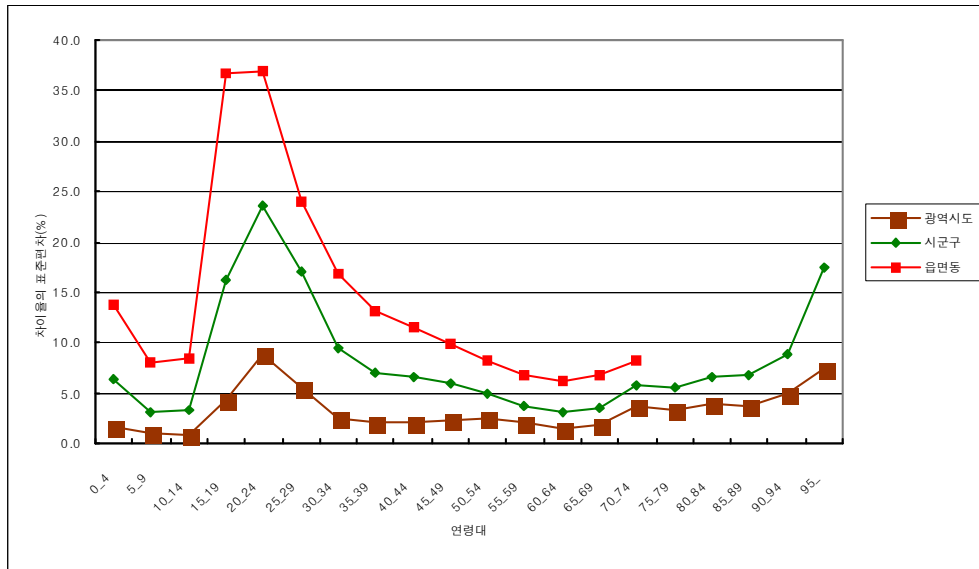
이주, 유학이나 취업에 따른 해외 거주 등의 이유로 인해서 가구면접조사의 형태로 이루어지는 인구주택총조사에서 조사되지 않았을 가능성이 크다. 이와 같은 해석은 이동성이 강한 연령층에서 인구주택총조사의 오차가 상대적으로 크게 나타난다는 이지연(2007b)의 연구 결과에 의해서 뒷받침된다. 40세 이후의 연령대에서는 차이율이 낮아지는데 이는 청년층의 높은 차이율을 발생시킨 여러 가지 원인들이 약화된 결과로 해석할 수 있다.

노년층에서는 상반된 방향의 차이율이 발생하였다. 부의 차이율이 발생한 원인으로는 과거 행정등록의 미비로 인해 주민등록집계에 잡히지 않는 인구층의 존재를 추정할 수 있다. 반면 정의 차이율이 발생한 원인은 사망 시 신원불명 등으로 인하여 주민등록이 정확하지 못하거나 복지시설 및 의료시설 수용로 인해 조사 응답여건이 안 되어 총조사에서 제외된 경우를 추정할 수 있다.

다음으로는 지역수준별로 차이율의 평균과 표준편차를 모아서 비교하였다. (그림 5-14)는 4가지 지역수준별로 연령대별 주민등록인구와 총조사인구의 차이율의 평균을 제시하였고 (그림 5-15)은 3가지 지역수준별로 연령대별 주민등록인구와 총조사인구의 차이율의 표준편차를 제시하였다.



(그림 5-14) 지역단위별 연령대별 차이율 평균 비교



(그림 5-15) 지역단위별 연령대별 차이율 표준편차 비교

연령대별 차이율 평균을 보면 연령대별 패턴 자체는 모든 지역단위 수준에서 유사하지만 규모는 지역수준별로 차이가 뚜렷했다. 이건(2007)의 결과와 유사하게 광역단위에서 소지역단위로 갈수록 차이율이 더 커졌다. 특히 광역시도와 시군구 수준 사이에서 차이율이 크게 차이났다.

대체로 전국 수준에서 차이율이 높은 연령대들에서 지역수준별 차이도 컸다. 반면 15-19세와 20-24세의 연령대는 이러한 경향의 예외로서 주목할 만하다. 전국 수준이나 광역시도에서는 차이율이 높지 않았는데 시군구, 읍면동 수준에서는 차이율이 상당히 높았다. 전국수준의 차이율과 소지역 수준의 차이율 사이에 나타난 격차는 인구이동시 발생한 주민등록과 실주거지 사이의 괴리에서 발생한 것으로 보인다. 이 경우 특정 소지역에서 나타난 차이가 다른 소지역의 차이와 서로 상쇄되면 광역수준에서는 차이율이 축소된다.

연령대별 차이율의 표준편차를 보면 소지역수준으로 갈수록 표준편차가 컸다. 특히 15-19세, 20-24세의 연령대에서 표준편차가 가장 컸다. 반면 전국 수준에서 차이율이 가장 높았던 35-39세, 30-34세의 연령대에서는 차이율의 표준편차가 상대적으로 적었다. 이러한 결과는 차이의 지역별 분포가 연령대에 따라서 다른 형태로 나타난다는 사실을 의미한다. 30대의 경우 주민등록과 총조사인구의 차이가 전국적으로 비교적 균일하게 높은 반면, 반면 10대

후반에서 20대 초반의 차이는 지역별로 큰 편차를 가졌다.

나. 시도 수준

<표 5-13>에서는 각 광역시도별로 차이율과 연령대 절대차이율 평균값을 제시하였다. 연령대 차이율의 산술평균값은 각 연령대간의 상쇄효과가 있을 경우 차이의 규모에 대한 정도를 정확히 반영할 수 없다. 이러한 문제를 해결하기 위해서 각 시도 연령대 차이율의 절대값을 평균한 연령대 절대차이율 평균을 제시하였다.

광역시도별로 비교해 보면 차이율이 높은 지역으로는 전라남도가 7.7%로 가장 높았고 다음으로 전라북도(5.7%), 제주도(4.8%), 충청남도(4.2%) 등이었다. 차이율이 낮은 지역으로는 광주광역시(-0.9%), 대전광역시(1.1%), 충청북도(1.8%) 등이 있었다. 서울특별시와 부산광역시는 각각 4.0%, 3.5%로서 평균적인 수준의 차이율을 보였다.

연령대 절대차이율 평균은 차이율과 상당히 다른 분포를 보였다. 전라남도처럼 차이율과 연령대 절대차이율 평균이 모두 높은 경우도 있었다. 하지만 광주광역시와 대전광역시의 경우 차이율은 전국에서 가장 낮은 수준이었으나 연령대 절대차이율 평균은 모두 4.5% 대 이상을 기록하여서 다른 지역과 거의 같은 수준이었다. 연령대별로 상이한 방향의 차이가 발생하여 서로 상쇄됨으로써 지역의 차이율이 낮게 나타난 것이었다. 이러한 사례들은 차이율이 낮다고 해서 주민등록의 활용가능성이 다른 시도보다 높다고 평가할 수 없음을 보여준다.

<표 5-13> 차이율 및 연령대 절대차이율 평균의 시도별 비교

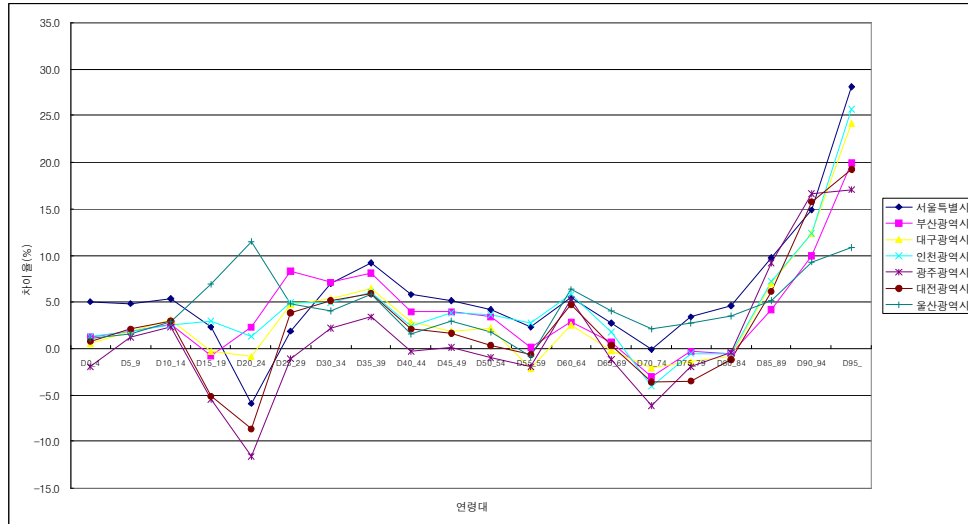
행정구역명	주민등록 인구수	총조사 인구수	차이율 (%)	연령대 절대차이율 평균 (%)
서울특별시	10,167,344	9,762,546	4.0	6.4
부산광역시	3,638,293	3,512,547	3.5	4.3
대구광역시	2,511,306	2,456,016	2.2	4.1
인천광역시	2,600,495	2,517,680	3.2	4.8
광주광역시	1,401,745	1,413,644	(0.9)	4.5
대전광역시	1,454,638	1,438,551	1.1	4.7
울산광역시	1,087,648	1,044,934	3.9	4.5
경기도	10,697,215	10,341,006	3.3	4.3
강원도	1,513,110	1,460,770	3.5	6.1
충청북도	1,488,803	1,453,872	2.4	4.2
충청남도	1,962,646	1,879,417	4.2	5.0
전라북도	1,885,335	1,778,879	5.7	6.2
전라남도	1,967,205	1,815,174	7.7	8.1
경상북도	2,688,491	2,594,719	3.5	5.2
경상남도	3,160,431	3,040,993	3.8	4.5
제주도	557,569	530,686	4.8	4.4

(그림 5-16)과 (그림 5-17)을 보면 약간의 예외들을 제외하고 대체로는 광역시는 광역시끼리, 도는 도끼리 서로 유사한 연령대 차이율의 유형을 가진다. 광역시와 도 사이에 특히 두드러진 차이를 보이는 연령대는 15-29세와 70-84세의 연령대였다.

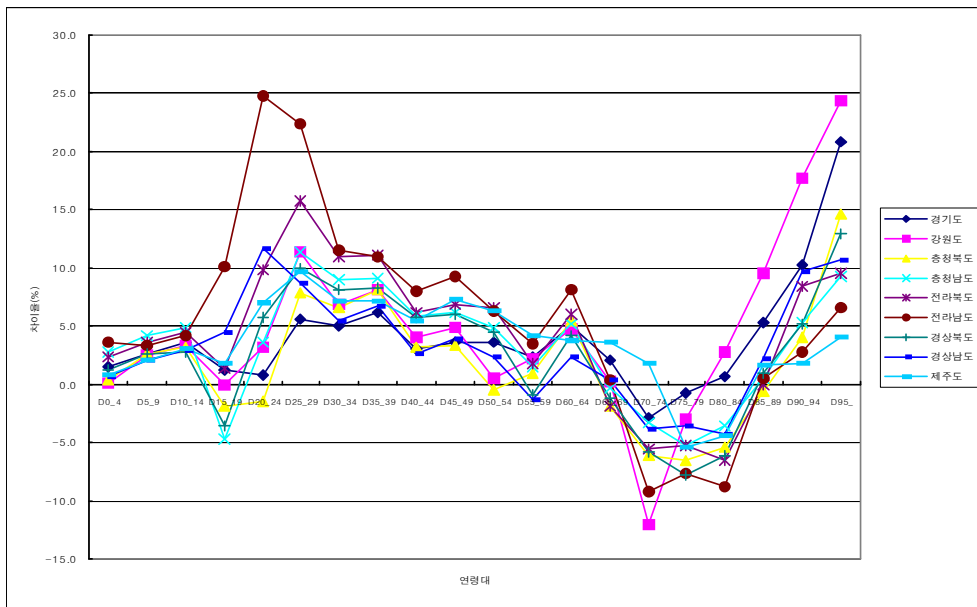
15-29세의 경우 광역시에서는 울산을 제외하고 부의 차이율을 보이거나 차이율의 절대값이 그리 크지 않았던 반면, 도에서는 경기도를 제외하고 매우 높은 정의 차이율을 보였다. 광역시에서는 상대적으로 젊은 연령층이 주민등록보다 많고 도에서는 젊은 연령층이 주민등록보다 적은 것이다. 이러한 현상은 젊은 층이 농촌지역과 중소도시에서 대도시로 이동한 결과로 보인다.

70-84세의 경우 광역시에서는 차이율의 절대값이 그리 크지 않았으나 도에서는 대체로 매우 낮은 부의 차이율을 보였다. 이들 노년층에서 주민등록

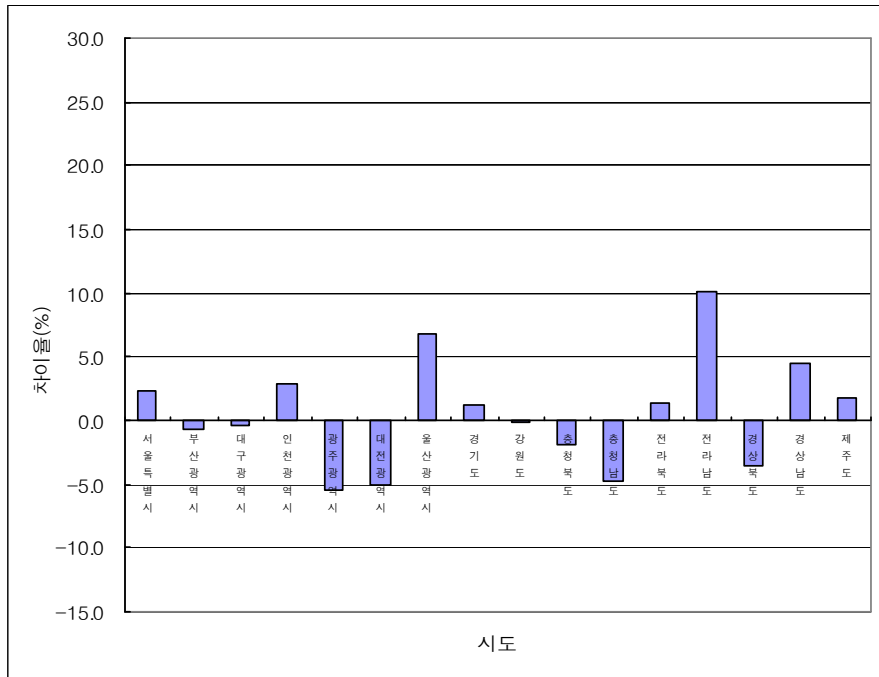
보다 총조사인구가 더 많은 것은 행정등록의 미비나 주민등록 관리 소홀에 따른 것이며 이러한 문제가 도시지역보다 농촌지역에서 더 심한 것이다. 상대적으로 차이율이 높았던 30대나 80대 중반 이후의 경우에는 시도 간 정도의 차이는 있지만 뚜렷한 유형의 차이는 나타나지 않았다.



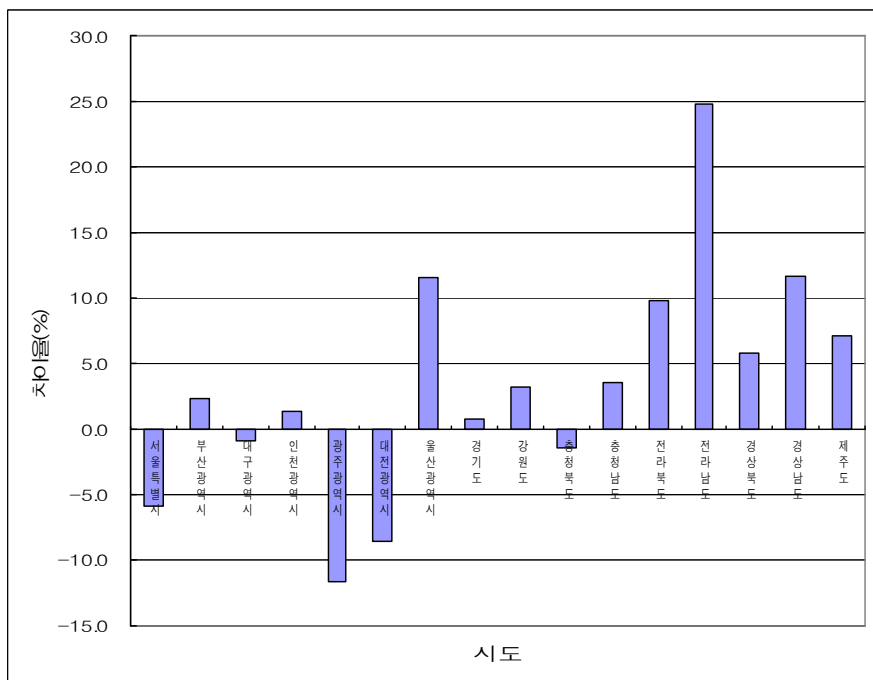
(그림 5-16) 연령대별 차이율 : 광역시 비교



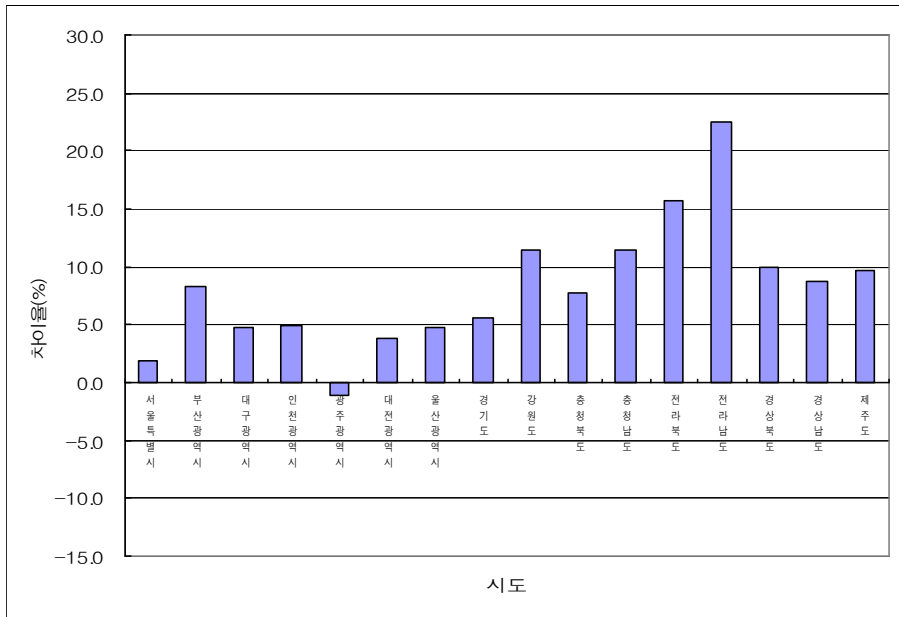
(그림 5-17) 연령대별 차이율 : 도 비교



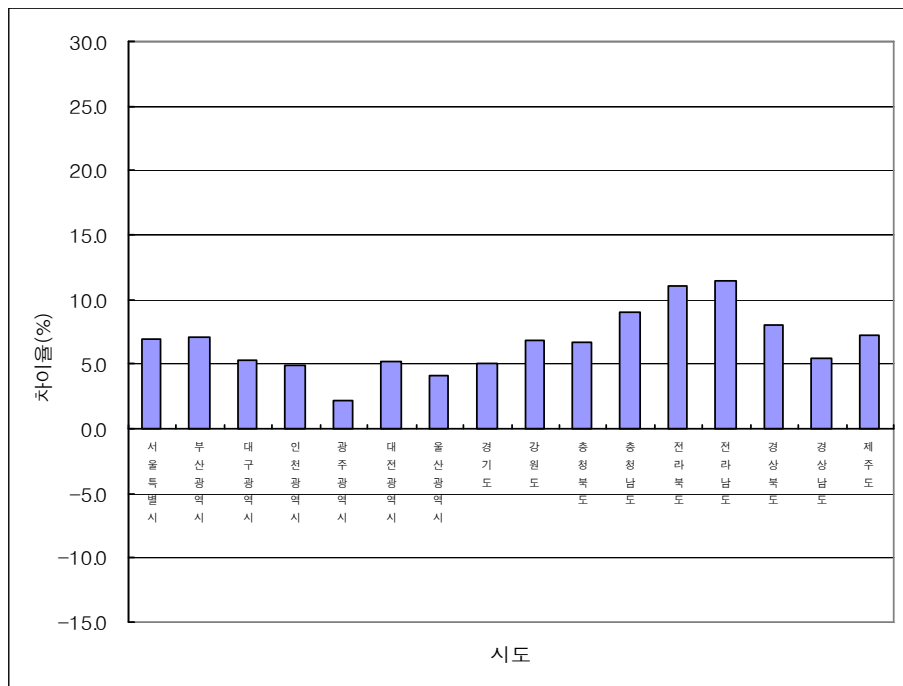
(그림 5-18) 15-19세 연령대 차이율



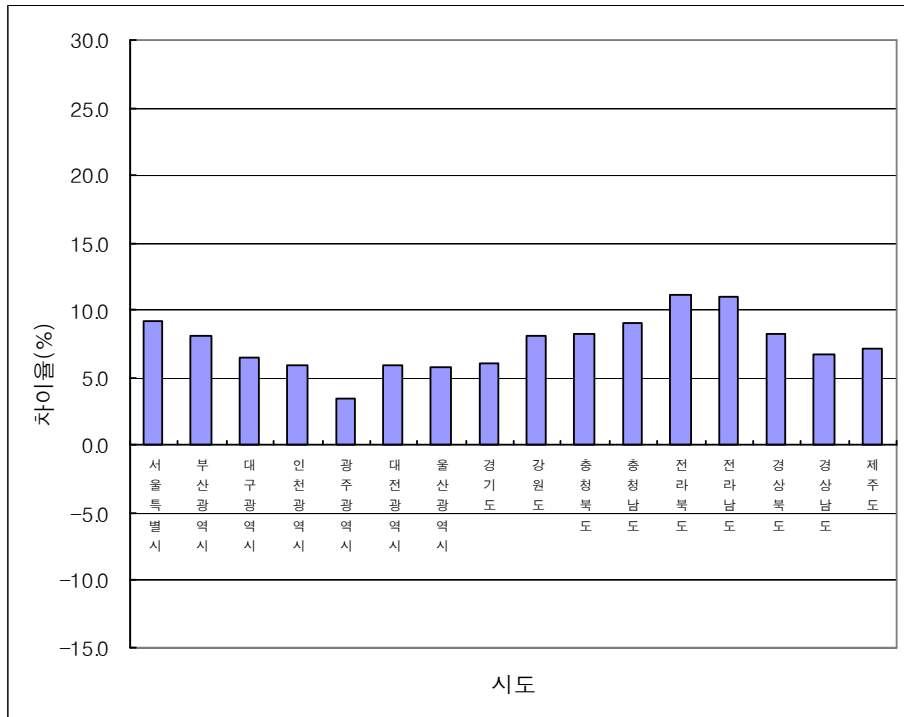
(그림 5-19) 20-24세 연령대 차이율



(그림 5-20) 25-29세 연령대 차이율



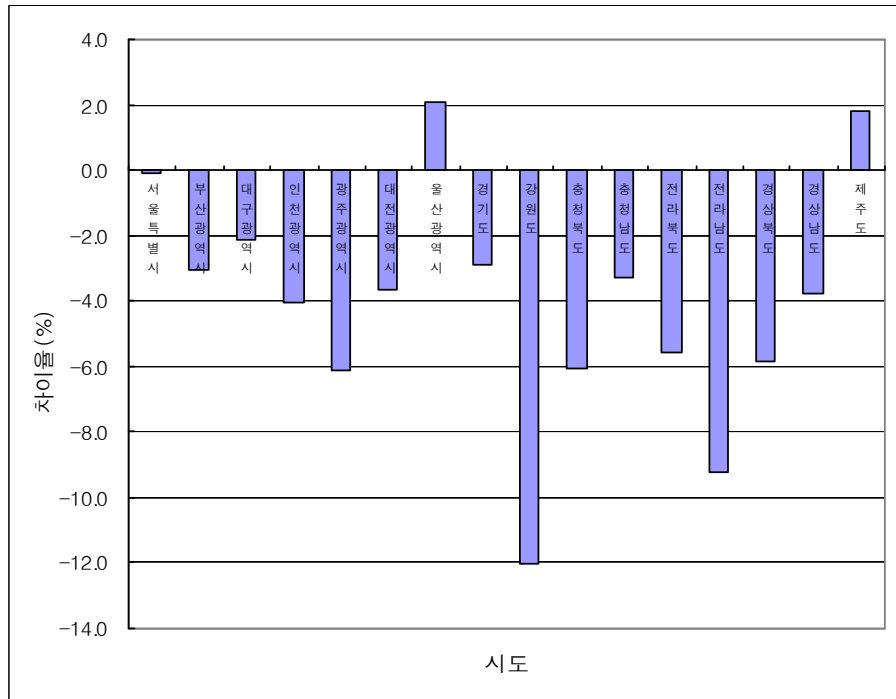
(그림 5-21) 30-34세 연령대 차이율



(그림 5-22) 35-39세 연령대 차이율

(그림 5-18)부터 (그림 5-22)까지는 15세 이상부터 30대까지의 여러 연령층에서 광역시도별 차이율을 비교한 결과이다. 이는 앞서 (그림 5-16)과 (그림 5-17)에서 제시한 시도 간 비교 결과를 연령대별로 세분하여 비교한 것이다. 분석의 대상이 된 연령대들 중에서 차이율의 시도 간 차이가 가장 두드러진 것은 20-24세 연령대였다. 차이율이 가장 높았던 지역은 전라남도였다. 반면 차이율이 최소값을 보인 지역은 광주광역시였다. 인접한 두 지역 사이에 서로 상반된 방향의 차이율이 나타나는 현상은 전라남도 출신 청년층이 광주로 이주한 결과로 추정된다. 마찬가지로 현상이 규모는 작지만 충청남도와 대전광역시의 사이에서 나타난다.

(그림 5-23)을 보면 70-74세 연령대에서는 대부분 총조사인구가 주민등록보다 많았고 대체로 광역시보다 도에서 차이율의 정도가 더 컸다. 농촌지역에서 주민등록의 정확성이 도시지역보다 더 미흡한 점을 고려할 때 이 연령대 차이율의 원인이 주민등록의 부정확성일 것이라는 추정을 뒷받침한다.



(그림 5-23) 70-74세 연령대 차이율

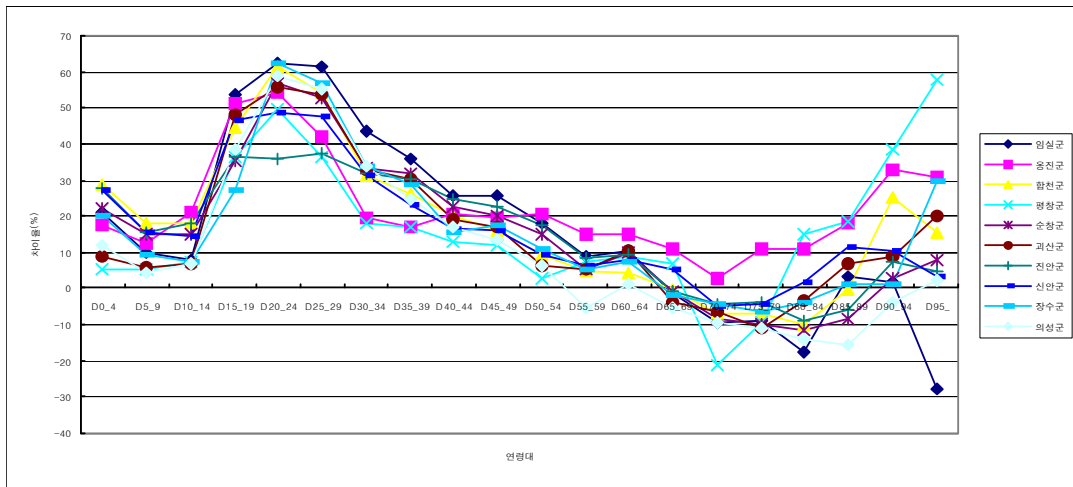
다. 시군구 수준

<표 5-14>는 시군구 수준에서 연령대 절대차이율 평균 기준 상위 10개 지역의 조사 결과이다. 연령대 절대차이율 평균이 주민등록의 활용 가능성을 보여주는 지표라는 점에서 상위 10개 지역은 주민등록을 통계자료로서 활용하기 가장 곤란한 지역이다. 연령대 절대차이율 평균이 가장 높은 지역은 전라북도 임실군이었고, 다음으로는 인천광역시 옹진군, 경상남도 합천군, 강원도 평창군, 전라북도 순창군 등의 순서였다. 상위 10개 시군구는 모두 농어촌지역이었고 전라북도 지역이 4개로 가장 많았다.

(그림 5-24)는 절대차이율 평균 기준 상위 10개 시군구의 연령대별 차이율을 제시하였다. 10개 시군구가 80대 이상의 고연령대를 제외하고는 대체로 유사한 형태였다. 특히 15세 이상 29세의 연령대에서는 모두 30%대 이상의 높은 차이율을 보였다. 이는 10대 후반과 20대가 주민등록자료를 통계자료로서 활용하기 가장 어려운 문제 연령대라는 사실을 재확인해 준다.

<표 5-14> 연령대 절대차이율 평균 기준 시군구별 차이율 : 상위 10개

시도	시군구	주민등록 인구	총조사 인구	차이율 (%)	연령대 절대차이율 평균(%)
전라북도	임실군	32,511	25,682	21.0	22.7
인천광역시	옹진군	15,609	12,261	21.4	22.2
경상남도	합천군	58,162	47,651	18.1	20.1
강원도	평창군	45,033	38,507	14.5	19.3
전라북도	순창군	32,012	26,250	18.0	19.2
충청북도	괴산군	38,595	32,208	16.5	18.0
전라북도	진안군	29,021	23,915	17.6	17.6
전라남도	신안군	46,451	38,429	17.3	17.5
전라북도	장수군	24,755	20,486	17.2	17.5
경상북도	의성군	64,457	56,522	12.3	17.0



(그림 5-24) 절대차이율 평균 기준 상위 10개 시군구별 연령대별 차이율

이와 같은 문제 연령대의 지역 분포에 대한 이해를 위하여 <표 5-15>에 서는 20-24세 연령대를 기준으로 총조사인구가 주민등록인구에 비해 적은 경우와 많은 경우 둘로 나누어 그러한 차이가 가장 큰 시군구 10개씩을 제시하였다.

<표 5-15> 20-24세 연령대 차이율 기준 시군구별 차이율

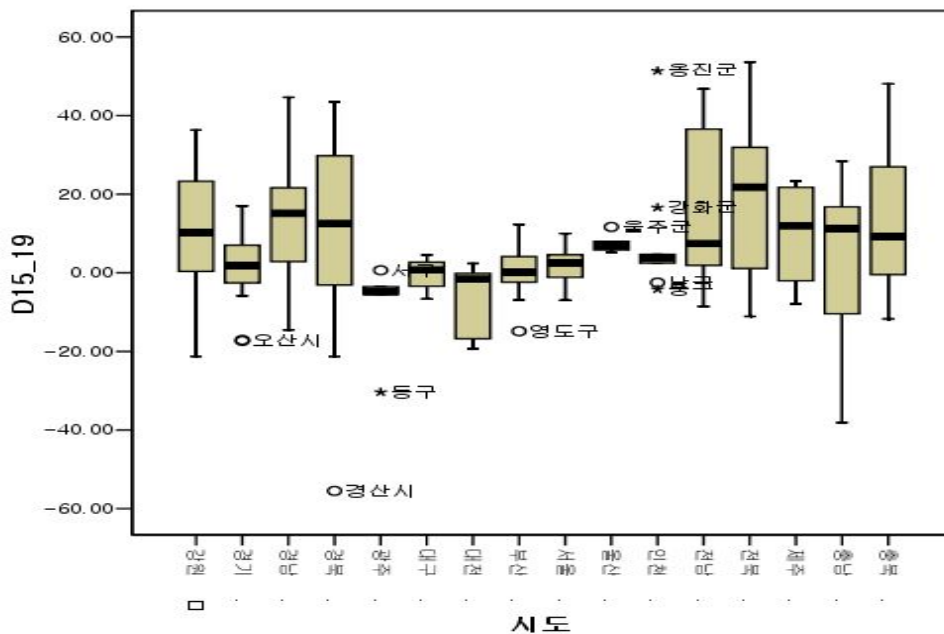
시도	시군구	10_14세	15_19세	20_24세	25_29세	30_34세	35_39세	70_74세
총조사인구가 적은 경우								
전라북도	장수군	7.4	27.4	62.4	56.6	33.7	28.5	(4.3)
전라북도	임실군	8.1	53.6	62.2	61.4	43.6	35.8	(9.6)
경상남도	합천군	17.8	44.7	61.5	54.3	31.3	26.2	(7.0)
경상북도	의성군	6.7	38.5	58.9	55.3	33.9	24.7	(9.5)
경상북도	영양군	5.3	15.3	57.5	51.7	29.4	19.1	(7.3)
전라남도	진도군	5.9	39.1	57.2	43.5	18.6	13.4	(12.3)
전라남도	장흥군	4.2	34.7	56.9	44.1	20.2	18.0	(8.9)
전라북도	순창군	14.9	35.3	56.9	52.6	33.5	31.7	(8.2)
전라북도	고창군	6.8	21.8	56.8	47.0	25.1	21.1	(10.4)
충청북도	괴산군	6.6	48.1	55.6	53.8	33.0	30.5	(6.3)
총조사인구가 많은 경우								
경상북도	경산시	(1.6)	(55.5)	(62.4)	(6.9)	4.3	5.9	(0.3)
광주광역시	동구	1.4	(30.4)	(40.0)	(7.3)	6.4	9.3	(9.3)
강원도	춘천시	3.1	(21.3)	(32.6)	(0.2)	4.8	5.5	(10.9)
충청남도	연기군	6.3	(22.1)	(30.7)	9.5	14.4	15.3	1.6
대전광역시	유성구	4.7	(19.3)	(30.1)	(4.5)	4.7	5.5	4.0
충청남도	아산시	5.7	(38.1)	(27.1)	(4.3)	5.5	8.8	1.3
경상북도	칠곡군	3.4	(18.4)	(26.0)	(12.4)	2.7	7.0	(1.7)
경기도	안성시	5.3	(17.1)	(22.9)	(0.4)	2.3	6.0	(6.8)
서울특별시	동작구	4.5	(6.5)	(21.4)	(8.9)	5.7	9.4	0.1
경기도	오산시	3.7	(17.3)	(21.3)	(3.7)	1.1	2.5	(5.0)

상위 10개 시군구 중에서는 전라북도가 4개 군이나 되어 가장 많았고 다음으로 전라남도와 경상북도가 각각 2개 군이었다. 대부분 농어업 위주의 산업구조를 가지고 있는 농어촌지역들이었다. 또한 20-24세 연령대의 차이율이 높은 지역은 인접한 15-19세와 25-29세 연령대에서도 차이율이 높았다. 이는 젊은 연령대의 높은 차이율이 특정 연령급에 한정된 현상이 아니라는 사실을 의미한다.

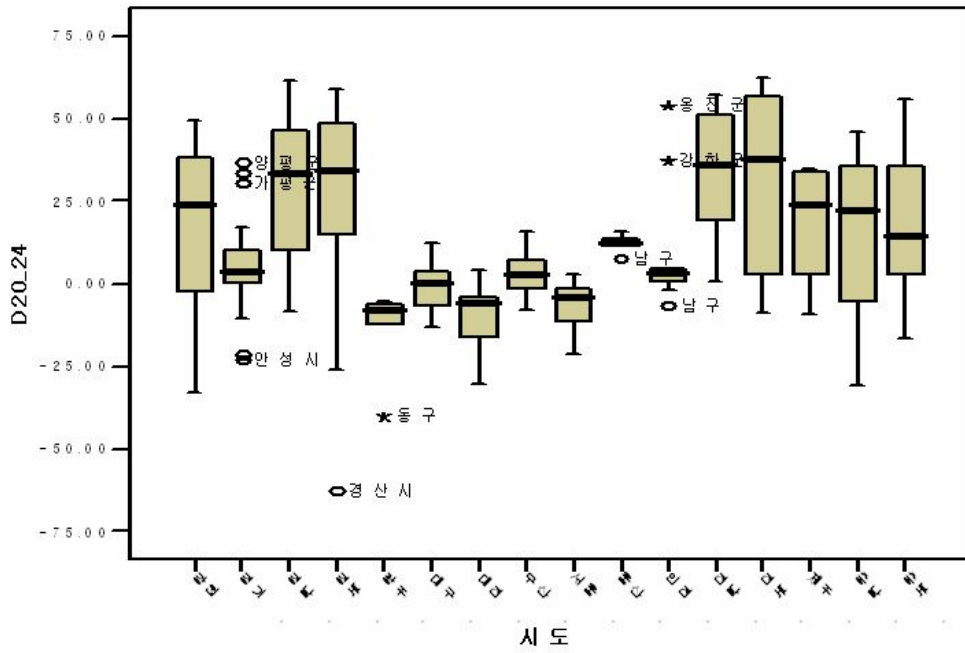
하위 10개 시군구들은 전국적으로 고르게 분포하였는데 대부분 대학이나

공단이 소재한 지역이었다. 이는 대학이나 공단에 유입된 젊은 층으로 인하여 주민등록과 인구주택총조사 사이에 큰 차이가 발생한 것임을 시사한다. 상위 10개와 하위 10개 시군구의 지역분포를 종합적으로 고려하면 젊은 층의 높은 차이율이 진학과 취업에 따른 인구이동에 의한 것으로 해석할 수 있다.

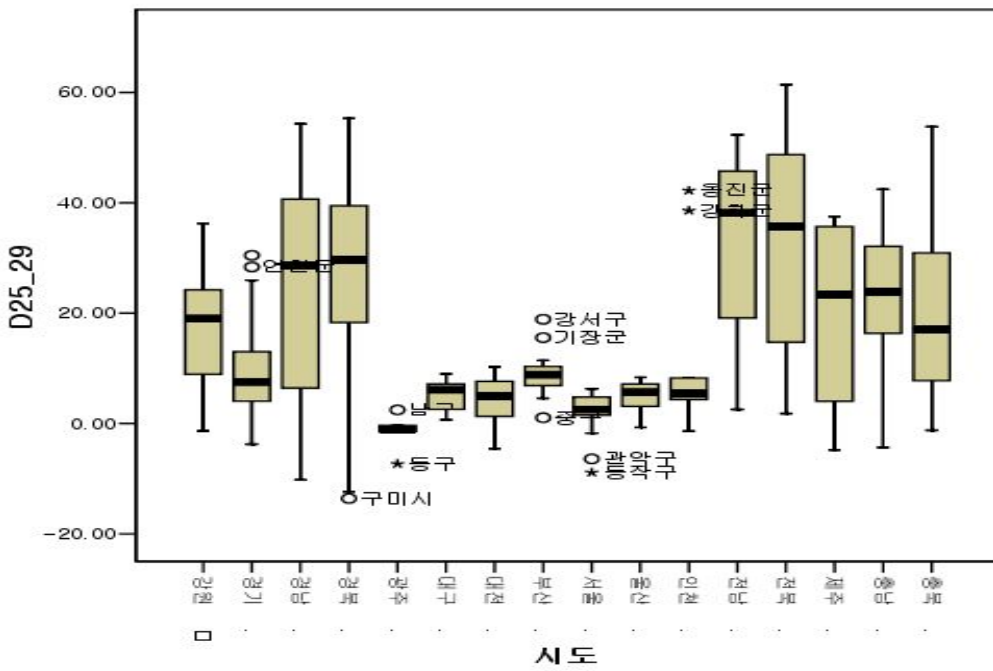
(그림 5-25)부터 (그림 5-30)까지 상자도표(box plot)를 사용하여 각 연령대별로 광역시도별 시군구의 차이율 분포를 제시하였다. 분포를 보면 대체로 광역시의 경우 시군구별 차이가 적은 반면 도의 경우 시군구별 차이가 컸다. 하지만 경기도의 경우 도의 유형이 아니라 광역시와 유사한 형태의 분포를 보였다. 이와 같은 사실은 경기도가 다른 도에 비해서 도시화와 산업화의 정도가 높은 현실과 연관되는 것으로 보인다.



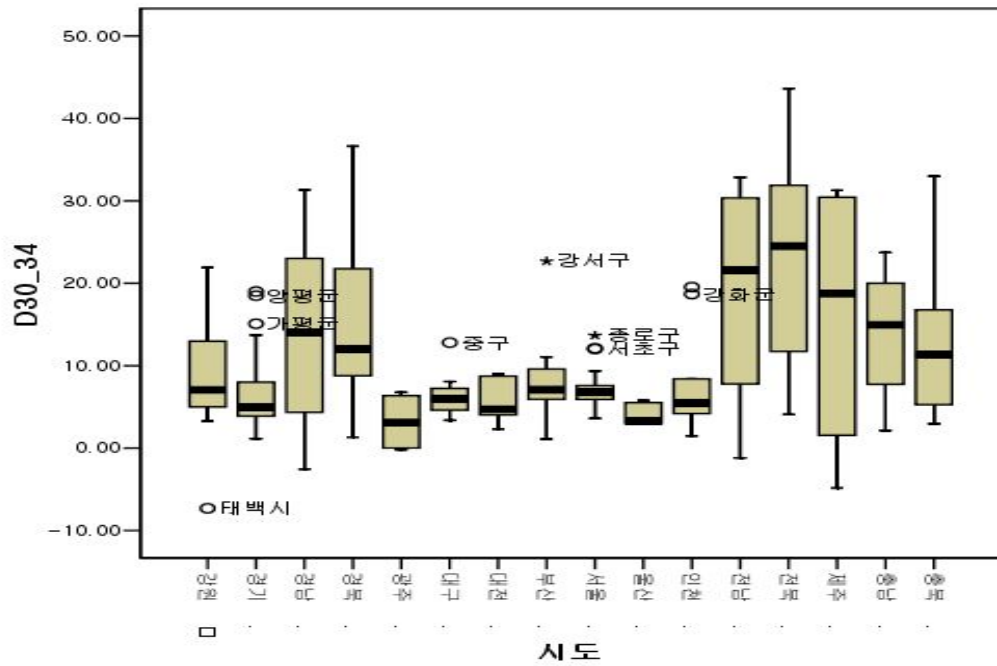
(그림 5-25) 15-19세 연령대 시군구별 차이율



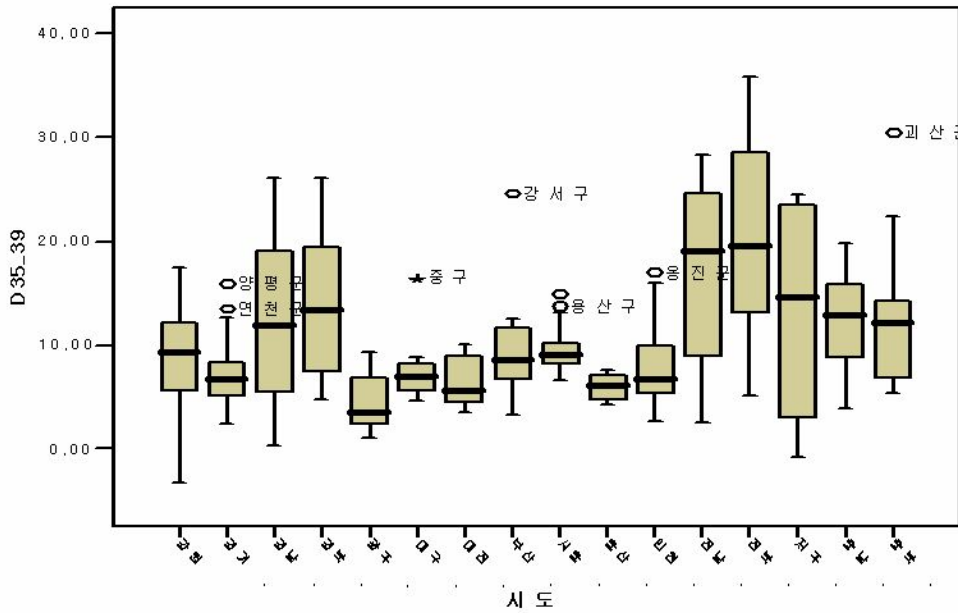
(그림 5-26) 20-24세 연령대 시군구별 차이율



(그림 5-27) 25-29세 연령대 시군구별 차이율



(그림 5-28) 30-34세 연령대시군구별 차이율



(그림 5-29) 35-39세 연령대 시군구별 차이율

포하였다.

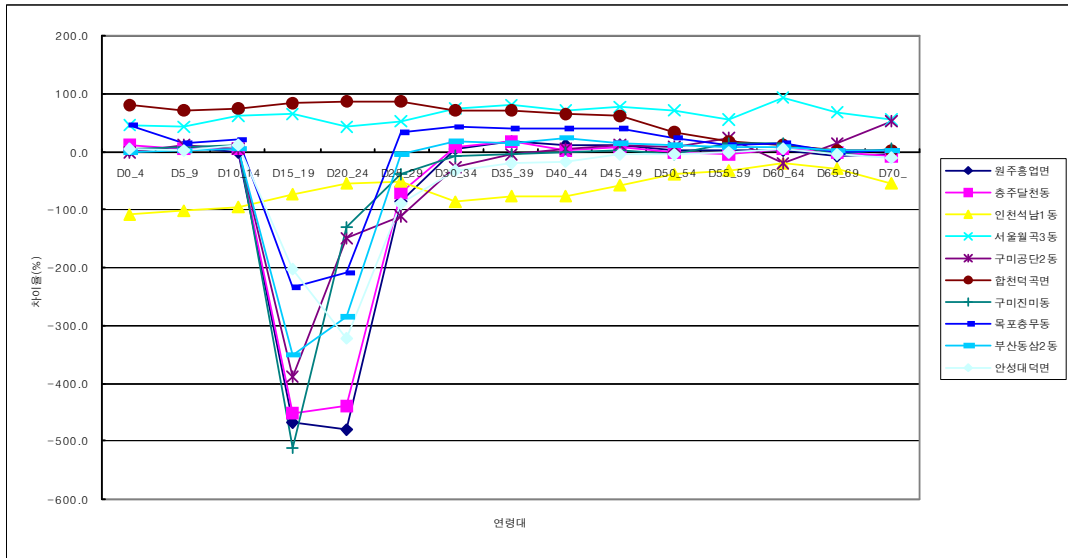
<표 5-16> 연령대 절대차이율 평균 기준 읍면동 차이율 : 상위 10개

시도	시군구	읍면동	주민등록 인구	총조사 인구	차이율 (%)	연령대 절대차이율 평균(%)
강원도	원주시	홍업면	6402	10703	(67.2)	73.7
충청북도	충주시	달천동	5124	8699	(69.8)	69.2
인천광역시	서구	석남1동	13913	23031	(65.5)	64.4
서울특별시	성북구	월곡3동	702	233	66.8	63.8
경상북도	구미시	공단2동	4491	7255	(61.5)	55.4
경상남도	합천군	덕곡면	1713	909	46.9	54.3
경상북도	구미시	진미동	15504	22954	(48.1)	51.2
전라남도	목포시	충무동	3744	4138	(10.5)	50.6
부산광역시	영도구	동삼2동	4972	6812	(37.0)	49.6
경기도	안성시	대덕면	8922	14175	(58.9)	49.1

연령대 절대차이율 평균이 높은 지역들의 특성을 살피기 위해서 (그림 5-31)에서는 상위 10개 읍면동 지역의 연령대별 차이율을 살펴보았다. 연령대의 차이율 유형은 크게 세 가지로 구별되었다. 10개의 읍면동 중 7개 지역은 15-19세, 20-24세의 연령대 차이율이 매우 낮은 경우였고, 2개 지역은 전체 연령대에서 차이율이 비교적 고르게 높은 경우였으며, 나머지 1개 지역은 전체 연령대에서 차이율이 비교적 고르게 낮은 경우였다.

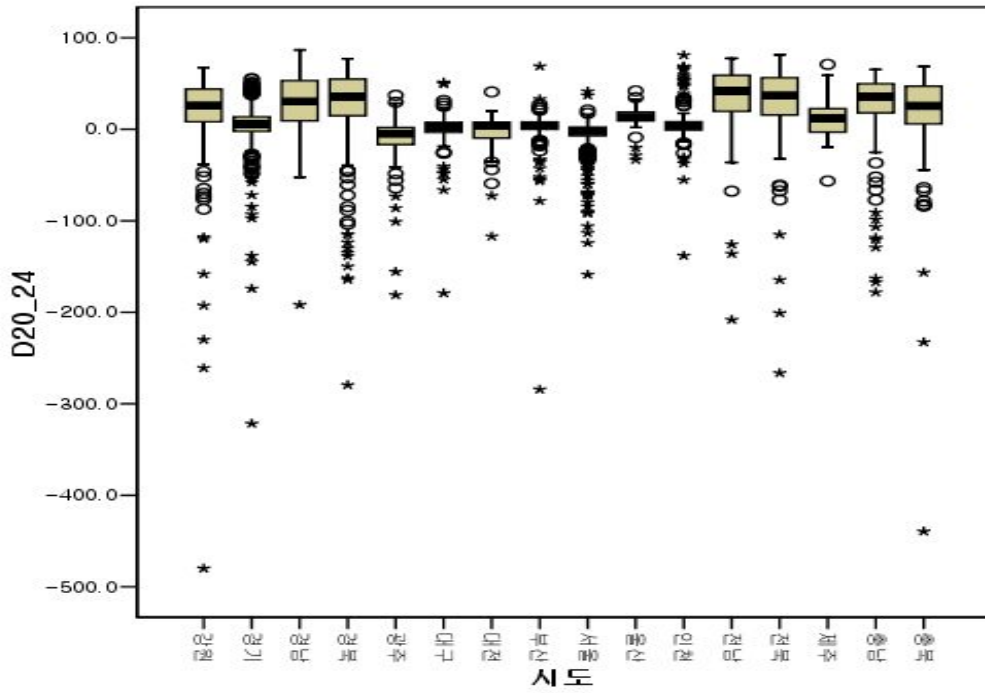
먼저 15-19세, 20-24세 연령대의 차이율이 매우 낮은 지역들을 보면 모두 대학 소재지와 공단 소재지의 특성을 갖고 있다. 원주시 홍업면, 충주시 달천동, 안성시 대덕면은 대학캠퍼스의 소재지로서 대학 타운이 형성되어 있다. 부산광역시 영도구 동삼2동과 목포시 충무동은 어촌 및 관광자원과 해양대라는 교육기관의 요소가 결합되어 있다. 구미시의 공단2동과 진미동은 구미시에 소재한 공단의 영향권에 속하였다. 다음으로 전체 연령대에서 차이율이 고르게 높았던 서울의 월곡동은 재개발 사업으로 주민등록인구가 실거주

자보다 전체 연령층에서 많은 것으로 보인다. 합천군 덕곡면의 경우는 전체적으로 개발이 부진하였다. 끝으로 전체 연령대에서 차이율이 고르게 낮았던 인천의 석남1동은 인근의 공업지역과 재개발이 관련된 것으로 보인다.

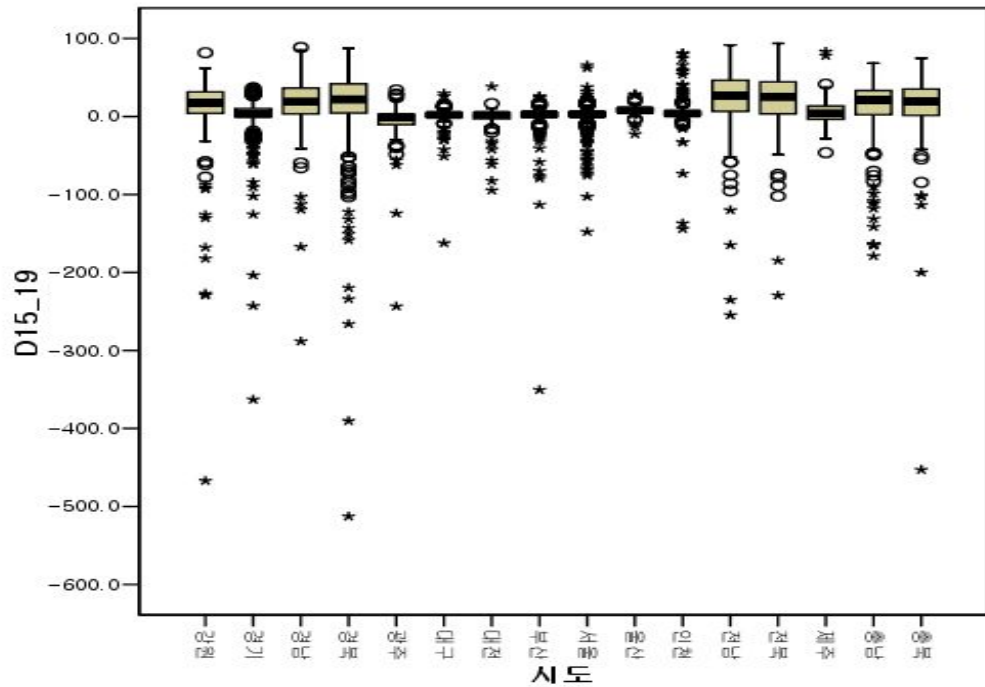


(그림 5-31) 절대차이율 평균 기준 상위 10개 읍면동 연령대별 차이율

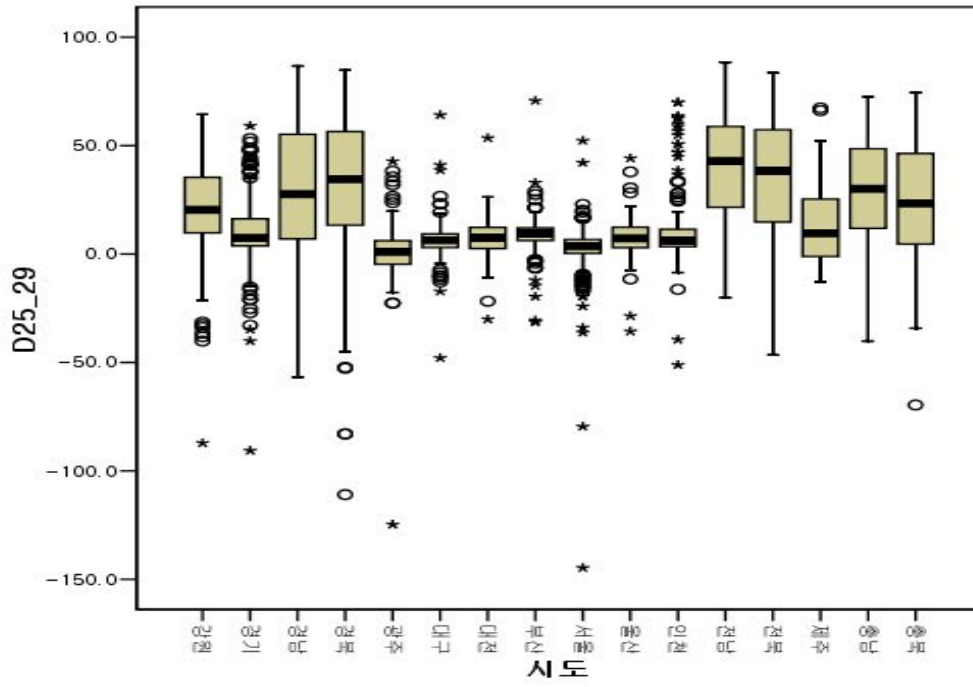
(그림 5-32)에서 (그림 5-37)은 문제가 되는 연령대별로 광역시도별 읍면동 차이율의 분포를 제시하였다. 이를 보면 읍면동별 차이율의 이상값이 가장 많이 나타난 연령대가 15-19세, 20-24세다. 이들 연령대에서는 중앙값에서 멀리 떨어져 있는 이상값들이 거의 모든 광역시도에서 나타났으며 그 정도도 매우 컸다. 특히 이상값들이 정보다 부의 차이율을 가지는 경우가 더 많았다. 15세-19세, 20-24세의 광역시도별 읍면동 차이율 분포를 보면 부의 방향으로 상당히 긴 꼬리를 가지고 중간값은 대체로 정의 방향에 위치해 있다. 긴 꼬리에 해당하는 지역들은 앞서 본 상위 10개 읍면동들처럼 젊은 연령층을 흡수하는 기관의 소재지로 보인다. 소수의 지역들이 젊은 층을 흡수하면서 높은 수준으로 부의 차이율을 보이는 반면 다수의 지역들은 젊은 층의 조사인구가 주민등록에 못 미치는 정의 차이율을 보였다. 이와 같이 상반된 방향의 차이율이 상쇄되는 까닭에 이 연령대의 읍면동 차이율이 상당한 규모임에도 불구하고 전국수준 차이율이 그리 높지 않았음을 알 수 있다.



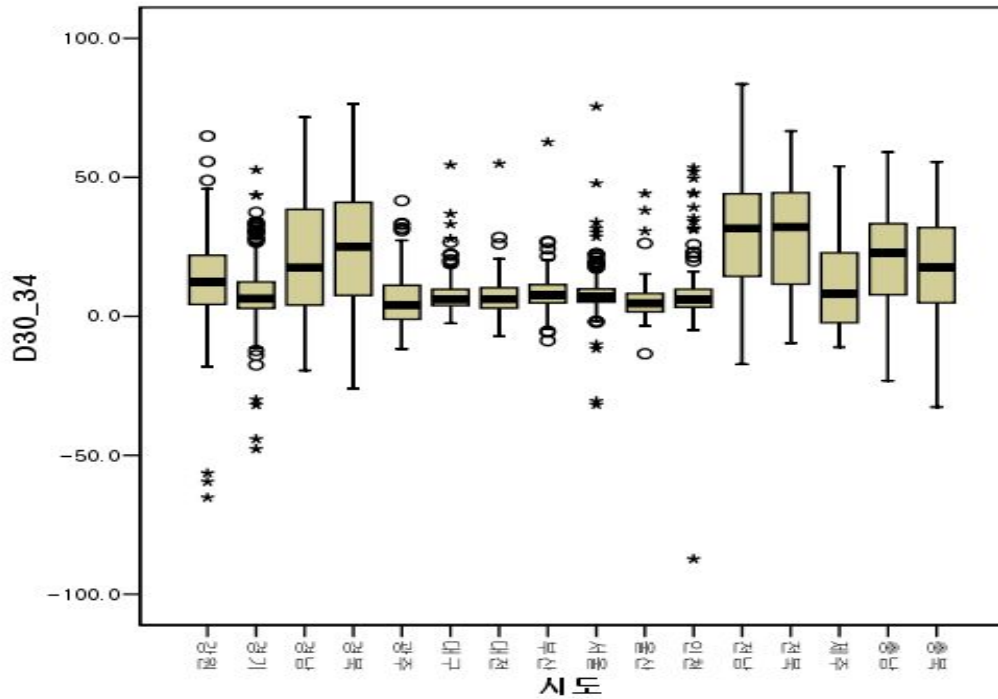
(그림 5-32) 20-24세 읍면동별 차이율



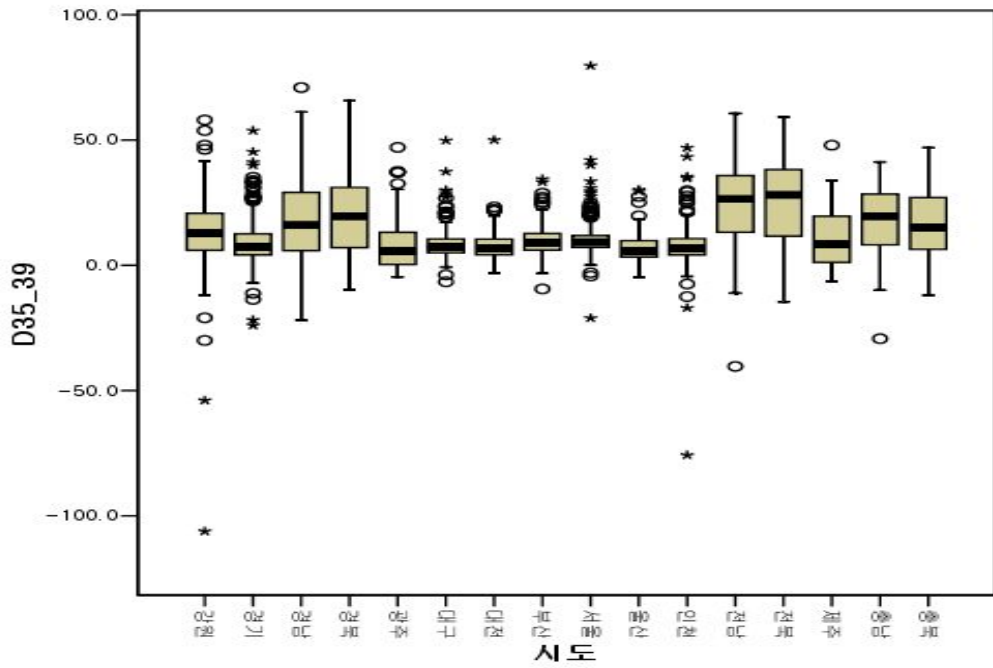
(그림 5-33) 15-19세 읍면동별 차이율



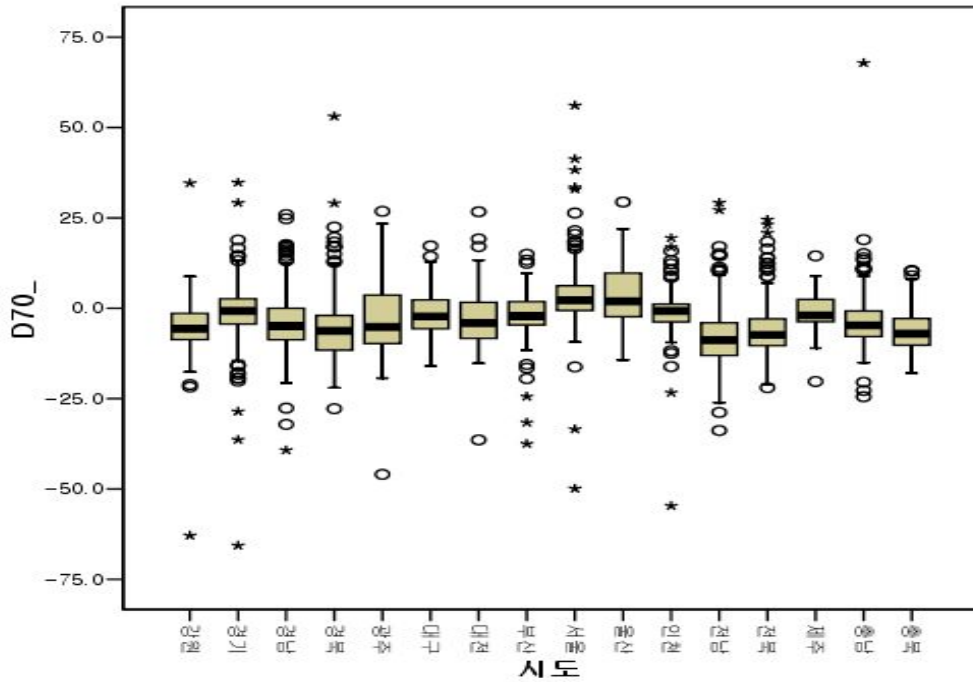
(그림 5-34) 25-29세 읍면동별 차이율



(그림 5-35) 30-34세 읍면동별 차이율



(그림 5-36) 35-39세 읍면동별 차이율



(그림 5-37) 70세 이상 읍면동별 차이율

3. 소결

이상의 분석은 연령대별로 차이율을 파악하는 것이 주민등록의 활용 가능성을 평가하는 데 매우 중요하다는 사실을 보여준다. 집계자료를 바탕으로 주민등록과 총조사인구의 차이를 파악하고자 할 경우 전국, 광역시도, 시군구, 읍면동의 수준에서 모두 연령대별 차이율을 세심하게 분석해야 한다. 연령대나 소지역을 고려하지 않고 특정지역에 대한 차이율을 계산할 경우 정의 차이율과 부의 차이율이 서로 상쇄됨으로써 차이율이 과소평가되는 위험이 발생한다.

또한 연령대 비교의 결과를 보면 주민등록인구와 총조사인구의 차이가 우연적으로 발생하는 것이 아니었다. 차이율이 연령대에 따른 일정한 유형을 보인다는 것은 주민등록인구와 총조사인구의 차이가 생애주기와 밀접한 관련이 있음을 보여준다. 노년층에서 나타나는 주민등록 미비처럼 과거 특정한 행정현실이 과도기적으로 작용한 경우도 있다. 하지만 이러한 사례들의 비중이 현재도 양적으로는 그리 크지 않고 앞으로는 점차 감소할 것으로 예상된다. 따라서 전체적으로는 생애주기가 주민등록과 총조사인구의 차이에 대한 가장 중요한 설명 요인이 된다. 경제적, 사회적인 이유로 활동이 활발해질수록 해외거주, 조사누락 등으로 총조사에 포함되지 않게 됨으로써 주민등록과의 차이가 커지는 경우들이 증가한다. 그리고 노년층에서는 조사누락이나 사망신고 미비 등 주민등록신고의 부정확으로 인해서 차이율이 다시 높아진다.

가장 주목할 만한 연령대가 15-24세의 연령층이다. 이 연령층에서는 대학진학이나 경제활동을 시작하면서 대학과 공단 소재지로 집중 이동하는 상황이 발생한다. 이때의 이동이 주민등록이전을 수반하지 않는 경우가 많아서 주민등록과 총조사인구 사이에 큰 차이가 나타나는 지역들이 존재함을 여러 지역수준에서 확인하였다.

이와 같이 주민등록인구와 총조사인구의 차이가 우연하게 나타나는 것이 아니라 구조적인 성격을 갖는다는 사실은 두 조사의 차이를 예측하고 설명할 수 있는 가능성을 의미한다. 또한 오차를 개선할 수 있는 가능성을 찾을 수 있다. 15-24세의 경우 대체로 젊은 층 유입 기관들에 의해서 오차가 발생

한다는 점을 고려하면, 대학의 학적이나 공단의 인사기록 활용을 통하여 주민등록 집계계의 한계를 개선할 수 있을 것으로 예상된다.

끝으로 연령대별 차이를 보면 주민등록의 부정확성에 따른 오차뿐만 아니라 인구주택총조사의 부정확성에 따른 오차의 가능성도 나타난다. 이와 같은 분석 결과는 주민등록을 통한 인구총조사의 대체가 아니라도 주민등록의 정보를 활용하여 인구주택총조사를 보완함으로써 오차를 줄이는 것이 현실적 대안이 될 수 있음을 시사한다.

제5절. 개인수준의 매칭

주민등록자료와 총조사자료의 차이를 명확히 이해하기 위해서는 최종적으로 두 자료를 개인 수준에서 1 대 1로 대응시켜 비교하여야 한다. 그러나 현재 그러한 자료는 존재하지 않는다. 총조사에서 개인정보에 관련된 자료를 입력하지 않고, 주민등록자료도 공개되지 않기 때문이다. 따라서 개인수준에서 두 자료를 완벽하게 1 대 1 대응시키기 위해서는 통계청에서 특별히 그러한 자료를 만들고 같은 시점에 행정자치부에서 주민등록자료를 얻어야 한다. 이 연구에서는 그 수준의 자료를 사용할 수 없었기 때문에, 정보가 모자란 가운데 개인을 식별할 수 있는 방법들을 점검해 보았다.

1. 자료

매칭대상 자료로는 2005년 인구주택총조사를 위해 2004년에 시행한 시범 조예행조사(이하 시범조사) 자료와 조사와 동일 시점의 2004년도를 주민등록 자료를 이용하였다. 먼저 각 자료의 특징을 살펴보면 다음과 같다.

가. 주민등록자료

제공받은 주민등록자료는 엑셀 파일로 되어 있다. 파일에는 주민등록자료는 ‘성명’, ‘성별’, ‘주소’, ‘생년월일’, ‘세대주관계’로 이루어져 있다. (그림

5-39)는 받은 주민등록자료 파일의 일부이다. 개인정보보호를 위해 이름은 삭제하였다.

엑셀에서는 앞자리수가 '0'이면 이를 '0'은 빼고 유효 숫자만 보여준다. 따라서 생년월일에서 1900년-1909년, 2000년-2005년생은 생년월일이 6자리가 아닌 3자리에서 5자리 숫자로 입력이 된다. 예를 들어서 2000년 05월 03일에 태어난 사람의 생년월일은 000503인데 엑셀에서는 503만 입력이 된다. 그런데 1900년 05월 03일 생도 주민등록의 앞자리는 똑같다. 따라서 현재 주어진 자료만으로는 100살 차이가 나는 1900년생과 2000년생을 구별할 수 없다.

실제 주민등록번호에서 주민등록번호의 7번째 자리는 성별을 구분(홀수 남자, 짝수 여자) 하지만, 1900년대에 태어난 사람과 2000년대에 태어난 사람을 구별하는 데 사용된다. 그러나 이 자료에는 주민등록번호의 앞 6자리(생년월일)만 있으므로 503이 2000년 05월 03일에 태어난 사람인지 1900년 05월 03일에 태어난 사람인지를 확인할 수 없다.

1	읍면동	성명	성별	주소						생년월일	세대주관계
2	내북면		여	충청북도	보은군	내북면	대안리	1	27	10620	손
3	내북면		남	충청북도	보은군	내북면	대안리	1	27	300928	본인
4	내북면		남	충청북도	보은군	내북면	대안리	1	27	570425	자
5	내북면		여	충청북도	보은군	내북면	대안리	1	27	750216	자부
6	내북면		여	충청북도	보은군	내북면	대안리	1	27	340910	처
7	내북면		남	충청북도	보은군	내북면	대안리	121	1 27	500303	본인
8	내북면		남	충청북도	보은군	내북면	대안리	121	1 27	770515	본인
9	내북면		여	충청북도	보은군	내북면	대안리	121	1 27	541118	모
10	내북면		여	충청북도	보은군	내북면	대안리	121	1 27	250917	조모
11	내북면		남	충청북도	보은군	내북면	대안리	129	27	290123	본인
12	내북면		남	충청북도	보은군	내북면	대안리	129	27	460323	본인
13	내북면		여	충청북도	보은군	내북면	대안리	129	27	520628	처
14	내북면		여	충청북도	보은군	내북면	대안리	129	27	240305	모
15	내북면		남	충청북도	보은군	내북면	대안리	129	27	830326	자

(그림 5-38) 주민등록자료의 형식

이렇게 정보가 모자란 상황에서, 세대주관계는 이를 구별하는 데 사용될 수 있다. 그렇지만 이것이 그다지 손쉬운 일은 아니다. '세대주관계'에는 상당히 많은 경우의 수가 있기 때문이다. <표 5-17>은 그러한 관계를 모두 나열한 것이다. 세대주관계에는 본인(세대주), 처, 자, 조부, 조모 등 모두 70가지 유형이 존재한다. 또한 세대주 자신의 특성을 알아야 그와의 관계에서 나이가 얼마인지를 짐작할 수 있다.

있다(<표 5-18> 참조).

<표 5-18> 시범조사 자료에 입력된 값

시도	가구원번호	출생월	출생양음력
시군구	가구원수	출생일	입력일시
읍면동	성별	교육1	수정일시
조사구	가구주관계	교육2	외국인가구원번호
거처번호	세는나이	혼인상태	국가
가구번호	출생연	출생띠	

시범자료에서 시도, 시군구, 읍면동, 조사구, 거처번호, 가구번호, 가구원번호를 조합하면 각 개인을 구별하는 고유한 코드 값을 만들 수 있다. 고유한 값을 만들면 이를 주민등록자료와 1 대 1로 대응시켜야 한다. 그렇지만 자료의 내역이 달라 대응이 쉽지 않다. 앞서 제시한 2000년생과 1900년생의 경우 세대주와의 관계로 나이의 진실값을 알 수 있었다. 이제 서로 같은 나이의 사람을 찾는 경우, 여러 사람이 나올 것이다. 이때 세대주와의 관계와 상응하는 가구주와의 관계를 비교하여 양쪽 자료에서 동일 인물을 판별하는 범위를 좁힐 수 있다. 그런데 여기 문제가 발생한다. 주민등록의 세대주와의 관계는 70가지인데 반해 가구주와의 관계는 14가지만 있어 상응하는 값을 찾을 수 없기 때문이다. 따라서 양쪽 자료에서 동일한 조건을 찾을 수 없어 1 대 1로 대응시킬 수가 없다. 참고로 시범조사자료에서 가구주와의 관계는 <표 5-19>에 실었다.

한편, 연결 값으로 주소를 생각해 볼 수 있다. 이 또한 문제가 있다. 주민등록자료에서는 주소가 번지까지 나와 있지만 시범조사 자료에서는 읍면동까지만 나와 있기 때문이다. 이런 문제에 봉착하여 이제 양쪽 자료를 연계하여 동일 인물은 양 자료에서 1 대 1 대응이 불가능하고 동일인 일 것 같은 사람들이 얼마나 되는지, 분명히 동일인일 수밖에 없는 사람은 얼마나 되는지 등, 서로 대응이 가능한 부분(portion)의 양을 찾는 방향으로 연구를 진행하게 되었다.

<표 5-19> 가구주와의 관계

가구주	배우자의 부모	형제자매의 자녀·배우자
가구주의 배우자	손자녀·그 배우자	부모의 형제·그 배우자
자녀	증손자녀·그 배우자	기타 친인척
자녀의 배우자	조부모	기타 동거인
가구주의 부모	형제자매·그 배우자	

다. 자료의 크기

본 연구에 사용된 자료는 충청북도 보은군의 11개 읍면동 자료와 부산광역시 해운대구 우동의 자료를 사용하였으며 각 읍면동별 총인구는 아래의 <표 5-20>과 같다. 부산광역시 해운대구의 경우 주민등록자료는 우1동, 우2동으로 나누어지지 않고 우동으로 합쳐져 있다.

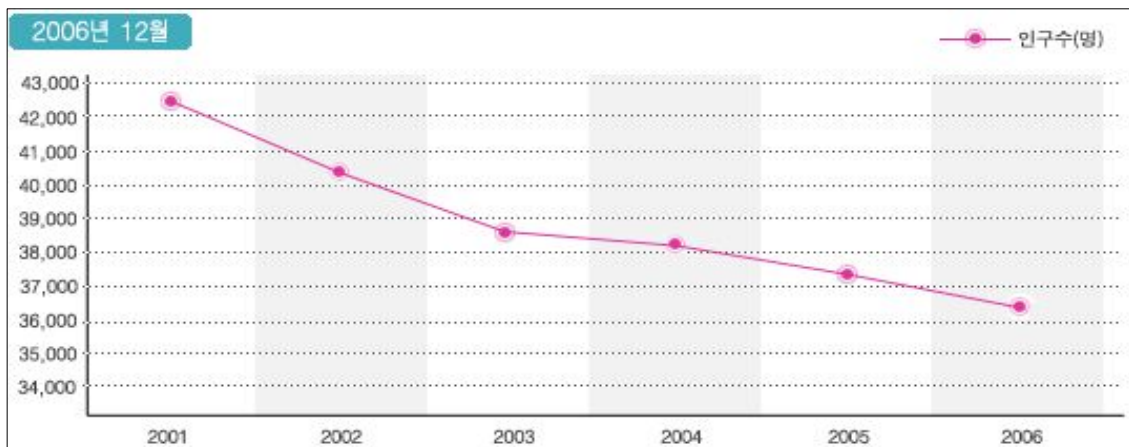
<표 5-20> 시도(시군구)별 총인구

시도(시군구)	읍면동	주민등록자료	시범조사 자료
충청북도 (보은군)	내북면	2,601	1,356
	산외면	2,846	1,456
	회북면	2,917	1,711
	보은읍	18,602	12,409
	내속리면	3,117	1,893
	회남면	1,076	484
	수한면	2,896	1,584
	외속리면	1,958	1,055
	탄부면	3,110	1,774
	삼승면	3,794	2,247
	마로면	4,051	2,146
합계		46,968	28,115
부산광역시 (해운대구)	우1동	50,000	29,418
	우2동		22,094
합계		50,000	51,512

그런데 충청북도 보은군의 주민등록자료 총인구가 기존에 공표된 인구수

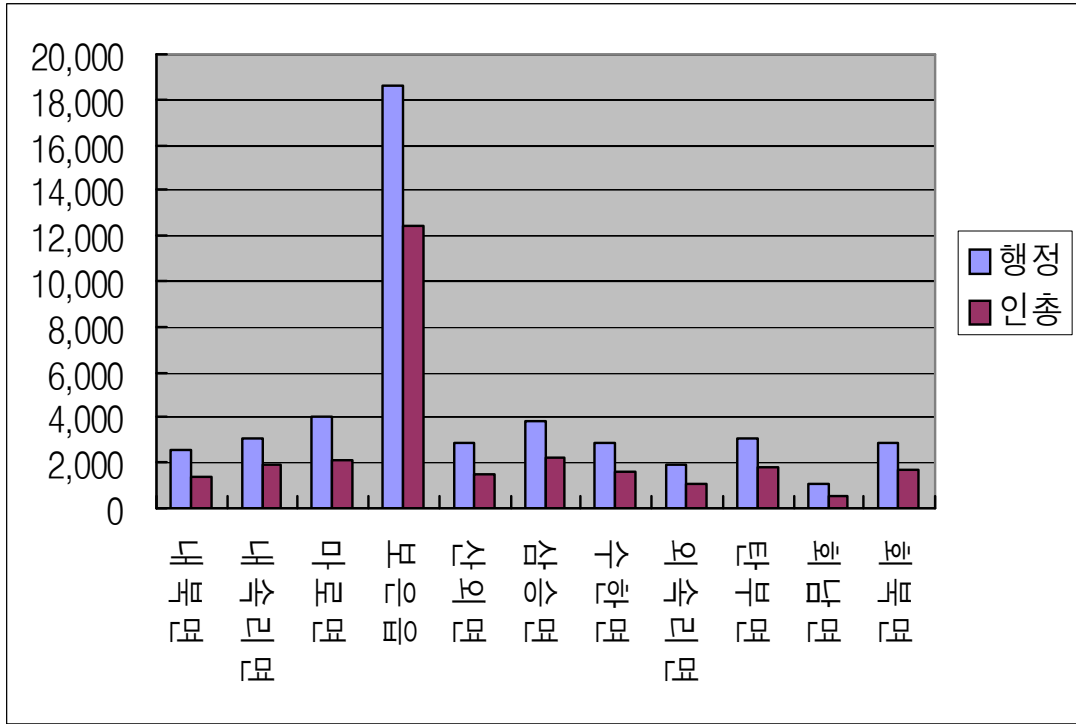
와 크게 차이가 나 문제가 되었다. 본 연구에 사용된 주민등록자료에서 충청북도 보은군의 인구를 보면 46,968명이다. 충청북도 보은군의 홈페이지에서 보은군의 인구 변화를 보면, (그림 5-40)과 같이 2006년에 36,293명이고 2004년 인구는 38,000에서 39,000명 사이에 있다. 또한, 이내성(2006)의 연구에서도 이 지역 인구는 34,374명임이 확인되었다.

이런 결과는 본 연구에서 제공받은 주민등록자료에는 오류가 있다는 것을 보여준다. 다행히 연구수행 마지막 단계에서 2005년 총조사 실제 자료와 2005년도 주민등록자료의 매칭을 시도한 통계개발원의 이내성(2006) 보고서가 발표되어 본 연구의 결과가 이내성(2006) 보고서의 참고 자료로서 기여할 수 있으리라 판단이 되어 연구결과를 정리하였다. 이런 측면에서, 자료문제로 말미하마 이 연구는 이내성(2006) 연구에 보완적인 연구가 되었다. 그렇지만 이 연구를 통해 두 자료를 연계하는 방법, 같은 사람을 찾아내는 알고리즘에 대한 상호 점검 및 정보교환을 할 수 있어 이 연구가 자료 연계방법을 개발하는 데 기여할 수 있을 것이다.



(그림 5-40) 충청북도 보은군의 연도별 인구 변화
출처 : 보은군 홈페이지

참고로 충청북도 보은군의 읍면동별 인구를 (그림 5-41) 제시하였다.



(그림 5-41) 충청북도 보은군의 인구

<표 5-21>은 충청북도 보은군의 남자와 여자의 인구를 나타낸 것이다. 주민등록자료에서는 남자가 52%, 여자가 48%이고, 시범조사 자료에서는 남자가 48%, 여자가 52%로 비율이 달라지는 것을 알 수 있다. 이내성(2006)에서 제시된 인구를 마지막 행에 제시하였는데, 2005년 시범조사 결과가 2004년 시범조사 때보다 남자는 0.7%포인트 많고, 여자는 0.7%포인트 적다.

<표 5-21> 성별 인구

	주민등록 인구		시범조사 인구	
	남자	여자	남자	여자
내북면	1,363	1,238	626	730
내속리면	1,648	1,469	905	988
마로면	2,042	2,009	980	1,166
보은읍	9,404	9,198	6,022	6,387
산외면	1,566	1,280	693	763
삼승면	2,001	1,793	1021	1,226
수한면	1,519	1,377	737	847
외속리면	1,030	928	497	558
탄부면	1,610	1,500	850	924
회남면	574	502	243	241
회북면	1,533	1,384	809	902
총합계	24,290 (51.7%)	22,678 (48.3%)	13,383 (47.6%)	14,732 (52.4%)
이내성(2006)	17,132 (49.8%)	17,242 (50.2%)	15,605 (48.3%)	16,715 (51.7%)

2. 매칭 방법

주민등록자료와 시범조사 자료를 매칭하기 위해서는 두 자료 모두 사람들을 구분할 있는 고유값이 존재해야 하는데 시범조사 자료의 경우에는 시도, 시군구, 읍면동, 조사구, 거쳐번호, 가구번호, 가구원번호를 조합하여 각 개인을 구분하는 고유값을 만들 수 있고, 주민등록자료의 경우에는 성명, 성별, 주소, 생년월일, 세대주 관계를 이용해서 고유값을 만들 수 있다.

두 개의 자료를 연결하기 위해서는 공통으로 입력되어 있는 값이 필요한데 성명은 시범조사 자료에는 입력되어 있지 않고, 주소의 경우 시범조사 자료에 읍면동까지의 정보만 입력되어 있으므로 읍면동까지 만을 사용할 수 있다. 또한, 시범조사 자료의 생년월일은 양력, 음력이 구분 가능하지만 주민등록자료는 구분이 불가능하므로 시범조사 자료의 생년월일이 양력으로 입력되어 있는 경우 음력으로 입력되어 있는 경우, 평달, 윤달 구분이 불가능하므로 모두 변환은 양력으로 변환을 한다.

본 연구에서 시범조사와 주민등록자료를 매칭하는 데 2가지 방법을 사용하였으며, 매칭을 위하여 사전에 정의한 내용이나 변환규칙은 다음과 같다.

가. 사전조정

- ① 주민등록자료에 대한 신뢰성이 떨어지므로 시범조사 자료를 기준으로 주민등록자료를 비교한다.
- ② 생년월일, 주소(시도, 시군구, 읍면동까지의 주소), 성별이 모두 같은 경우 시범조사에서 조사된 사람과 주민등록자료에 기록되어 있는 사람은 같은 사람으로 한다(주민등록자료에는 생년월일이 양력, 음력이 구분되어 있지 않지만 시범조사 자료에는 양력, 음력이 구분되어 있다).
 - 생년월일을 비교할 때 자료에 입력되어 있는 그대로 비교를 한다.
 - 시범조사 자료에서 양력으로 입력되어 있는 경우 음력으로 변환을 해서 주민등록자료와 비교를 한다. 예를 들어서 시범조사 자료에서 1977년 10월 2일(양력)로 생년월일이 입력되어 있는 경우 주민등록자료에서 1977년 10월 2일이 생일인 사람을 모두 찾은 후 1977년 10월 2일(양력)을 음력으로 변환을 하면 1977년 8월 20일이 되는데 주민등록자료에서 1977년 8월 20일이 생일인 사람을 모두 찾는 방법을 사용한다.
- ③ 시범조사 자료에서 음력으로 입력되어 있는 경우 양력으로 변환을 해서 주민등록자료와 비교를 한다(음력생일에서 태어날 월이 윤달이 있는 경우 평달, 윤달 2가지 모두 양력으로 변환해서 주민등록자료와 비교하는 방법을 사용하였다). 예를 들어서 시범조사 자료에서 1976년 8월 2일(음력)로 생년월일이 입력되어 있는 경우 주민등록자료에서 1976년 8월 2일이 생일인 사람을 모두 찾은 후 1976년 8월은 윤달이 있으므로 1976년 8월 2일을 평달로 가정하고 양력으로 변환을 해서 1976년 8월 26일이 생일인 사람을 모두 찾고, 1976년 8월 2일을 윤달로 가정해서 양력으로 변환을 해서 1976년 9월 25일이 생일인 사람을 모두 찾는 방법을 사용한다.

나. 시범조사자료를 기준으로 주민등록자료를 이용하는 방법(방법1)

시범조사 자료를 기준으로 주민등록자료에서 생년월일, 성별, 주소가 같은 사람을 모두 찾은 후 가장 먼저 나타나는 사람을 선택하는 방법이다. 구체적으로는 생년월일, 성별, 주소가 같은 사람을 모두 찾은 후 이 사람들을 주소별로 나누고 같은 주소에 2명이상 속해 있는 사람들만을 대상으로 가장 먼저 나타나는 사람을 선택하는 방법이다.

예를 들어서 시범조사 자료에서 부산시 해운대구 우동에 살고 있는 1976년 7월 9일(양력)에 태어난 남자가 있다고 한다면 주민등록자료에서 부산시 해운대구 우동에 살고 있으면서 1976년 7월 9일(양력)에 태어난 남자와 1976년 7월 9일을 음력으로 변환해서 1976년 6월 13일(음력)에 태어난 남자를 모두 찾는다. 찾은 결과 ‘갑’, ‘을’, ‘병’ 순서대로 3명의 정보가 같은 경우 ‘갑’만을 선택하고 ‘을’, ‘병’을 선택하지 않는 방법이다.

다. 주민등록자료의 주소를 이용하는 방법(방법2)

(방법1)은 ‘갑’, ‘을’, ‘병’ 중 제일 먼저 나온 사람을 매칭하는 것이었지만, (방법2)는 ‘갑’, ‘을’, ‘병’을 주소별로 구별하여(주민등록자료의 경우 번지수까지 나와 있으므로 ‘갑’, ‘을’, ‘병’의 주소가 모두 다를 수 있다) 2명 이상이 매칭된 것만을 정확하게 매칭되었다고 보는 방법이다. 예를 들면, 각각의 주소가 A, B, C이고 A주소에는 1명, B주소에는 2명, C주소에는 3명이 있을 경우 A주소에 있는 ‘갑’은 같은 주소에 1명만 있으므로 제외하고 ‘을’과 ‘병’ 중에서 ‘을’이 먼저 나타나므로 ‘을’을 선택하는 방법을 사용한다.

이때 관계를 고려할 여지는 있으나 두 자료의 관계 표현이 너무 다르기 때문에 효용성이 없다. (방법2)를 제안하게 된 이유는 주민등록자료의 주소별로 나누었을 때 같은 주소에 한명만이 존재를 하면 이 사람이 시범조사 자료에 있는 사람인지 우연히 생년월일과 성별, 읍면동까지의 주소만 같은 것인지 알 수 없으므로 좀 더 정확히 하기 위해서 같은 주소에 2명이상 있는 경우만을 따로 생각하게 되었다.

3. 매칭 결과

가. 충청북도 보은군

(1) 방법1

방법1을 이용하여 충청북도 보은군의 매칭 결과는 다음과 같다. 보은읍의 경우 시범조사 자료에서의 인구는 12,409명이고 주민등록자료는 18,602명이다. 시범조사 자료의 12,409명 중에서 주민등록자료와 생년월일, 성별, 주소가 같은 사람이 6,760명이고 54.5%(결과 / 시범조사)에 해당한다.

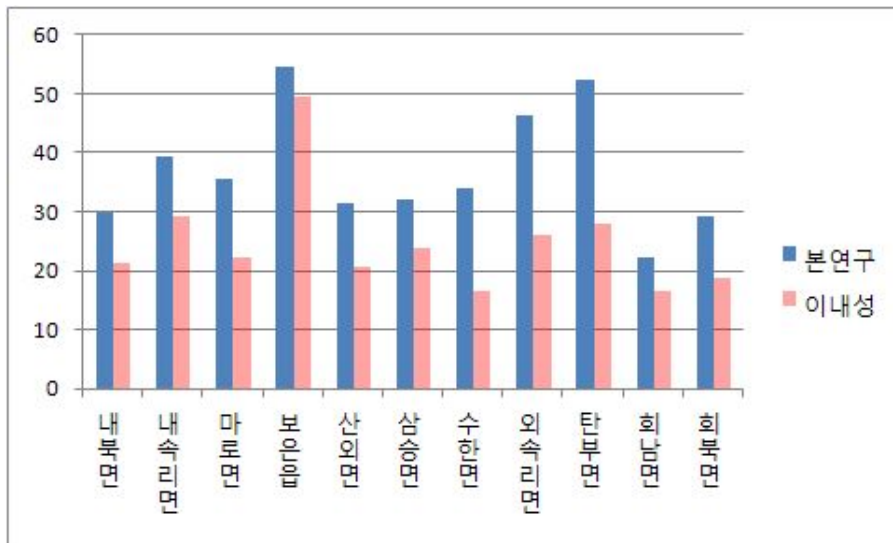
<표 5-22> 방법1을 이용한 매칭 결과

읍면동	주민등록 인구	시범조사 인구	일치 결과	일치율 (%)	이내성(2006)	
					일치 결과	일치 비율
내북면	2,601	1,356	405	29.9	347	21.3
내속리면	3,117	1,893	746	39.4	615	29.3
마로면	4,051	2,146	763	35.6	573	22.2
보은읍	18,602	12,409	6,760	54.5	6,960	49.6
산외면	2,846	1,456	455	31.3	358	20.7
삼승면	3,794	2,247	722	32.1	580	23.7
수한면	2,896	1,584	540	34.1	301	16.5
외속리면	1,958	1,055	489	46.4	351	25.9
탄부면	3,110	1,774	930	52.4	586	28.0
회남면	1,076	484	107	22.1	108	16.5
회북면	2,917	1,711	502	29.3	348	18.6
총인구	46,968	2,8115	12,419	44.2	11,127	34.4

이내성(2006)의 일치율 결과와 약 10% 차이가 나는 것으로 나타났다. 이러한 결과에 대해 본 연구에서 사용된 주민등록자료의 정확성이 의심스럽기 때문에 명확히 왜 그러한지를 제시할 수는 없다. 단, 이러한 자료에서는 매

칭방법이 정확한 매칭을 보장하는 것이 아니어서, 이것은 매칭률의 상한을 나타내기 때문에 30-40% 이하의 매칭률이라고 말할 수 있다.

한편, 두 연구 모두 인구의 약 40%가 거주하는 보은읍에서 50% 정도의 일치율을 보여 다른 면들과는 차이가 많이 났다. 이에 대한 분석은 두 개의 주민등록자료 비교가 필요한데 자료의 제약으로 할 수가 없었다.



(그림 5-42) 방법1을 이용한 매칭의 일치율 비교

성별, 연령별로 좀더 세분화해서 매칭되는 결과를 보면 다음과 같다.

<표 5-23> 방법1을 이용한 성별 매칭 결과

읍면동	남 자				여 자			
	주민등록 인구	시범조사 인구	일치 결과	일치율 (%)	주민등록 인구	시범조사 인구	일치 결과	일치율 (%)
내북면	1,363	626	220	35.1	1,238	730	185	25.3
내속리면	1,648	905	401	44.3	1,469	988	345	34.9
마로면	2,042	980	390	39.8	2,009	1,166	373	32.0
보은읍	9,404	6,022	3,405	56.5	9,198	6,387	3,355	52.5
산외면	1,566	693	247	35.6	1,280	763	208	27.3
삼승면	2,001	1,021	374	36.6	1,793	1,226	348	28.4
수한면	1,519	737	279	37.9	1,377	847	261	30.8
외속리면	1,030	497	256	51.5	928	558	233	41.8
탄부면	1,610	850	484	56.9	1,500	924	446	48.3
회남면	574	243	63	25.9	502	241	44	18.3
회북면	1,533	809	266	32.9	1,384	902	236	26.2
총합계	24,290	13,383	6,385	47.7	22,678	14,732	6,034	41.0

다음은 지역별, 연령대별로 나누어서 본 것이다.

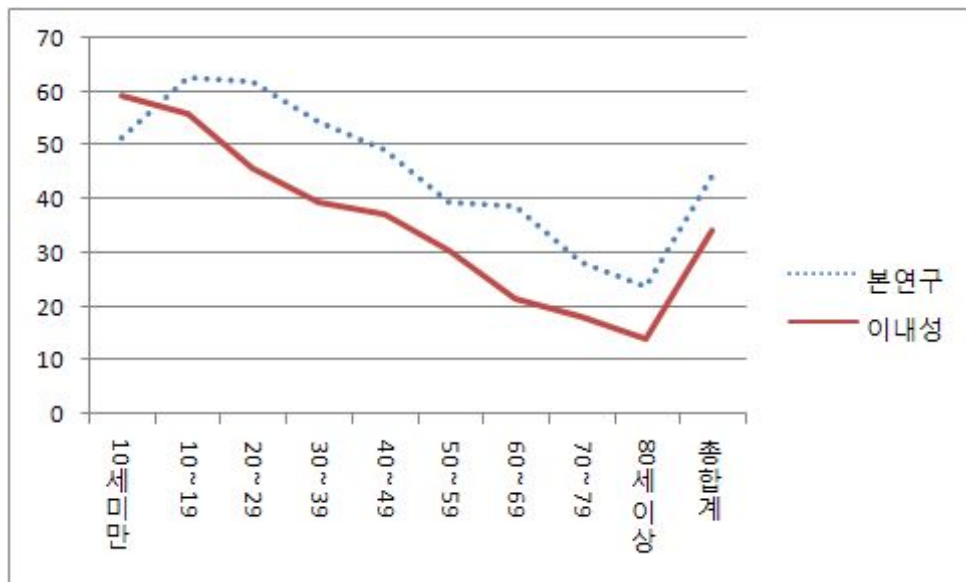
<표 5-24> 방법1을 이용한 연령별 매칭 결과

읍면동		10세미 만	10- 19	20- 29	30- 39	40- 49	50- 59	60- 69	70- 79	80- 89	90- 99	합계
내북면	일치수	31	49	17	22	63	62	79	66	14	2	405
	총수	78	95	39	83	167	183	300	294	102	14	1356
내속리면	일치수	69	124	49	72	121	123	117	62	8	1	746
	총수	116	217	97	161	291	343	358	236	71	3	1893
마로면	일치수	56	119	43	44	109	130	143	96	20	3	763
	총수	113	226	70	110	281	353	453	394	129	15	2146
보은읍	일치수	784	1176	599	1050	1202	734	688	405	111	11	6760
	총수	1483	1766	870	1693	2072	1556	1394	1116	376	43	12409
산외면	일치수	49	62	24	34	62	52	96	61	12	3	455
	총수	94	104	51	91	190	202	336	292	84	12	1456
삼승면	일치수	45	132	38	40	129	96	147	67	27	1	722
	총수	103	239	90	115	325	351	483	389	134	18	2247
수한면	일치수	28	60	33	32	73	75	96	98	38	7	540
	총수	74	108	72	96	172	242	334	329	136	21	1584
외속리면	일치수	25	74	30	38	65	67	109	63	15	3	489
	총수	55	110	54	76	130	162	220	174	65	9	1055
탄부면	일치수	33	93	62	54	119	141	226	145	52	5	930
	총수	76	144	98	85	191	279	393	352	139	17	1774
회남면	일치수	13	16	10	6	11	15	20	12	3	1	107
	총수	26	30	13	29	53	82	111	98	38	4	484
회북면	일치수	36	61	30	27	66	90	96	64	31	1	502
	총수	71	108	60	78	239	286	354	360	130	24	1711
합계	일치수	1169	1966	935	1419	2020	1585	1817	1139	331	38	12419
	총수	2289	3147	1514	2617	4111	4039	4736	4034	1404	180	28115

이내성(2006)과 비교하기 위하여 비율로 조사해 본 바로는 <표 5-25>와 (그림 5-43)에서 볼 수 있듯이 10세 미만을 제외하고는 전 연령대에서 본 연구 매칭비율이 높지 나타난 것으로 보아 앞서와 같이 느슨한 결론을 내릴 수 있을 것 같다. 첫째 특징으로는 본 연구에서 10세 미만이 낮은 이유는 주민등록자료 정리 중에 00년생이 빠져나갔기 때문이라고 여겨진다. 두 번째 특징으로는 매칭률이 연령의 증가에 따라 낮아지는 것은 생년월일 등의 오류 가능성이 높아져서 노령 인구가 커지는 만큼 매칭되는 수가 늘지 않는다고 볼 수 있다.

<표 5-25> 방법1을 이용한 연령별 매칭 일치율

읍면동	10세 미만	10 - 19	20 - 29	30 - 39	40 - 49	50 - 59	60 - 69	70 - 79	80세 이상	합계
내북면	39.7	51.6	43.6	26.5	37.7	33.9	26.3	22.4	13.7	29.9
내속리면	59.5	57.1	50.5	44.7	41.6	35.9	32.7	26.3	11.3	39.4
마로면	49.6	52.7	61.4	40.0	38.8	36.8	31.6	24.4	15.5	35.6
보은읍	52.9	66.6	68.9	62.0	58.0	47.2	49.4	36.3	29.5	54.5
산외면	52.1	59.6	47.1	37.4	32.6	25.7	28.6	20.9	14.3	31.3
삼승면	43.7	55.2	42.2	34.8	39.7	27.4	30.4	17.2	20.1	32.1
수환면	37.8	55.6	45.8	33.3	42.4	31.0	28.7	29.8	27.9	34.1
외속리면	45.5	67.3	55.6	50.0	50.0	41.4	49.5	36.2	23.1	46.4
탄부면	43.4	64.6	63.3	63.5	62.3	50.5	57.5	41.2	37.4	52.4
회남면	50.0	53.3	76.9	20.7	20.8	18.3	18.0	12.2	7.9	22.1
회북면	50.7	56.5	50.0	34.6	27.6	31.5	27.1	17.8	23.8	29.3
총합계	51.1	62.5	61.8	54.2	49.1	39.2	38.4	28.2	23.6	44.2
이내성 (2006)	59.3	56.0	45.8	39.5	37.3	30.3	21.4	18.1	14.0	34.4



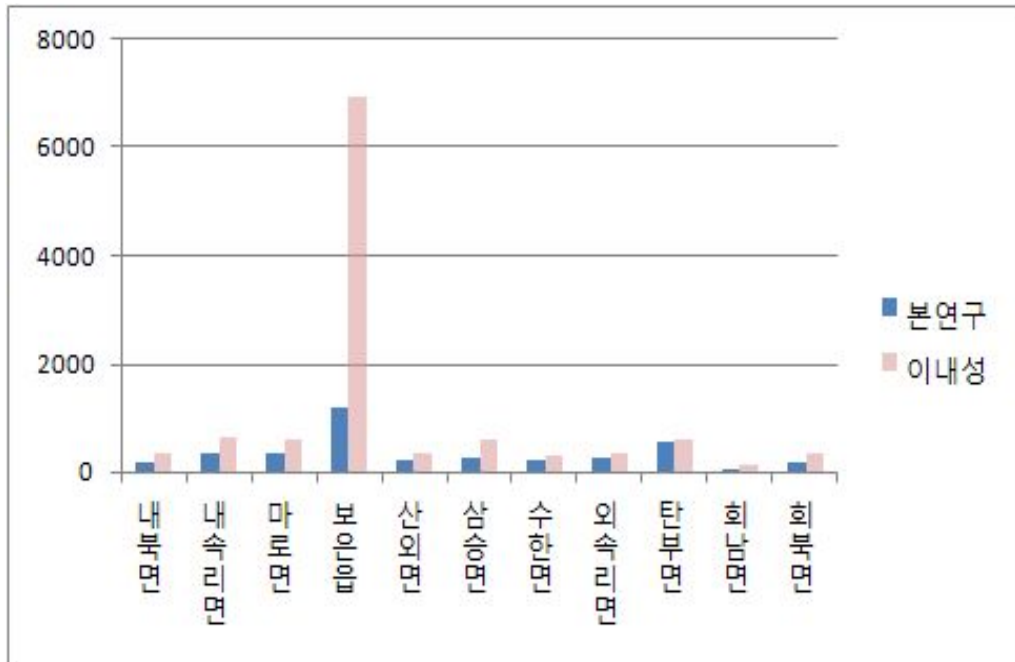
(그림 5-43) 방법1을 이용한 매칭의 일치율 비교

(2) 방법2

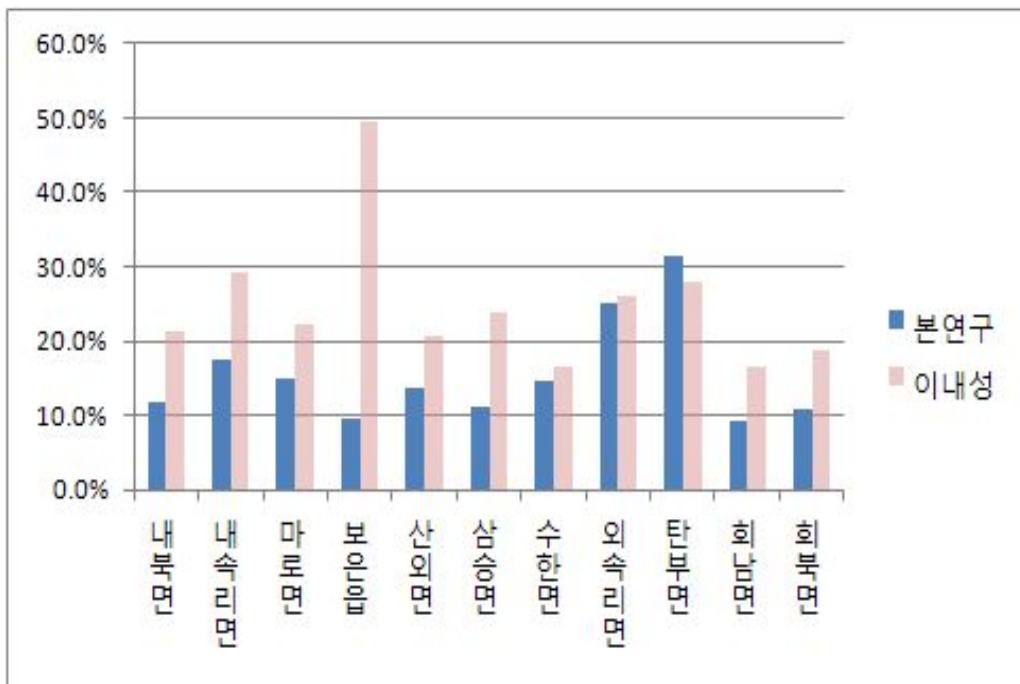
방법2를 이용하여 충청북도 보은군의 매칭 결과는 다음과 같다. 보은읍의 경우 시범조사 자료에서의 인구는 12,409명이고 주민등록자료는 18,602명이다. 시범조사 자료의 12,409명 중에서 주민등록자료와 생년월일, 성별, 주소가 같은 사람이 1,196명이고 이는 시범조사 총수의 9.6%에 해당한다. 방법1과 달리 이 값들은 매칭률의 하한으로 볼 수 있다. 왜냐하면 한 주소에 두 사람이상이 매칭되었을 때만 매칭된 것으로 본 것이기 때문에 거의 정확 매칭이라고 본다면 실제 매칭률은 13%보다는 높다고 할 수 있다. 방법1의 결과와 같은 관점으로 <표 5-26>, <표 5-27>, (그림 5-44), (그림 5-45)를 해석할 수 있다.

<표 5-26> 방법2를 이용한 매칭 결과

읍면동	주민등록 인구	시범조사 인구	일치 결과	일치율 (%)	이내성(2006)	
					일치 결과	일치율 (%)
내북면	2,601	1,356	159	11.7%	347	21.3
내속리면	3,117	1,893	333	17.6%	615	29.3
마로면	4,051	2,146	318	14.8%	573	22.2
보은읍	18,602	12,409	1,196	9.6%	6,960	49.6
산외면	2,846	1,456	198	13.6%	358	20.7
삼승면	3,794	2,247	248	11.0%	580	23.7
수한면	2,896	1,584	230	14.5%	301	16.5
외속리면	1,958	1,055	266	25.2%	351	25.9
탄부면	3,110	1,774	559	31.5%	586	28.0
회남면	1,076	484	45	9.3%	108	16.5
회북면	2,917	1,711	183	10.7%	348	18.6
총인구	46,968	28,115	3,735	13.3%	11,127	34.4



(그림 5-44) 방법2를 이용한 매칭의 일치수 비교



(그림 5-45) 방법2 이용한 매칭의 일치율 비교

성별, 연령별로 좀 더 세분화해서 매칭되는 결과를 보면 다음과 같다.

<표 5-27> 방법2를 이용한 성별 매칭 결과

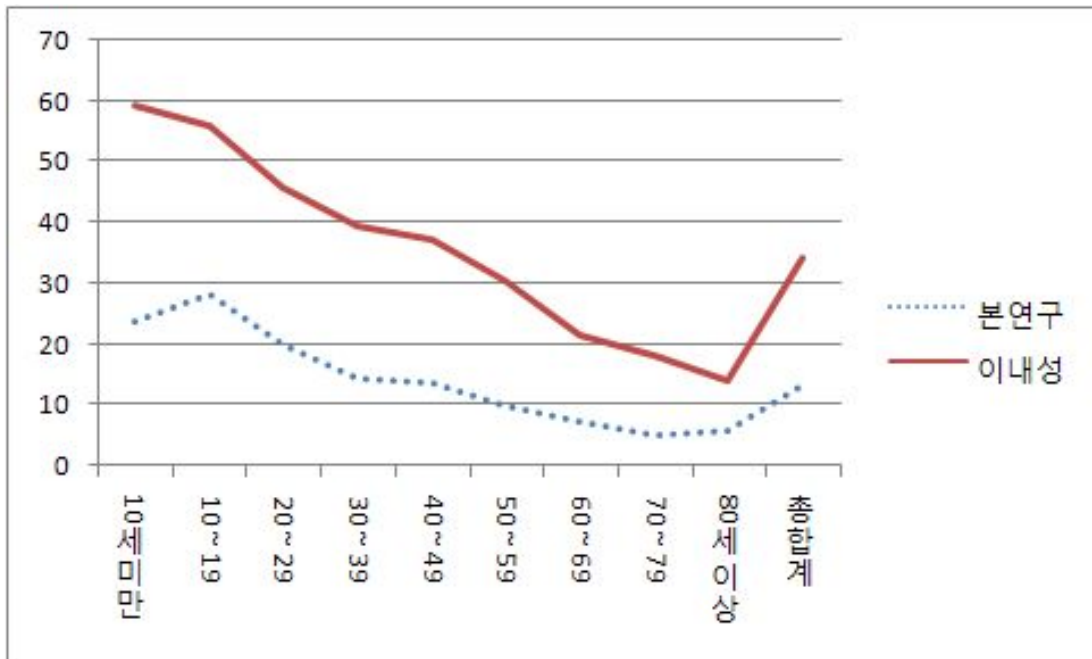
읍면동	남 자				여 자			
	주민등록 인구	시범조사 인구	일치결 과	일치 비율(%)	주민등록 인구	시범조사 인구	일치 결과	일치 비율(%)
내북면	1,363	626	75	12.0	1,238	730	84	11.5
내속리면	1,648	905	175	19.3	1,469	988	158	16.0
마로면	2,042	980	158	16.1	2,009	1,166	160	13.7
보은읍	9,404	6,022	604	10.0	9,198	6,387	592	9.3
산외면	1,566	693	101	14.6	1,280	763	97	12.7
삼승면	2,001	1,021	137	13.4	1,793	1,226	111	9.1
수한면	1,519	737	113	15.3	1,377	847	117	13.8
외속리면	1,030	497	137	27.6	928	558	129	23.1
탄부면	1,610	850	288	33.9	1,500	924	271	29.3
회남면	574	243	22	9.1	502	241	23	9.5
회북면	1,533	809	94	11.6	1,384	902	89	9.9
총합계	24,290	13,383	1,904	14.2	22,678	14,732	1,831	12.4

<표 5-28> 방법2를 이용한 연령별 매칭 결과

읍면동		10세	10-	20-	30-	40-	50-	60-	70-	80-	90-	합계
		미만	19	29	39	49	59	69	79	89	99	
내북면	일치수	24	37	7	9	21	24	16	14	6	1	159
	총수	78	95	39	83	167	183	300	294	102	14	1356
내속리면	일치수	52	87	24	47	49	36	27	10	1		333
	총수	116	217	97	161	291	343	358	236	71	3	1893
마로면	일치수	40	90	27	24	51	39	28	15	3	1	318
	총수	113	226	70	110	281	353	453	394	129	15	2146
보은읍	일치수	247	331	116	161	166	92	45	26	8	4	1196
	총수	1483	1766	870	1693	2072	1556	1394	1116	376	43	12409
산외면	일치수	35	44	12	21	34	16	22	10	1	3	198
	총수	94	104	51	91	190	202	336	292	84	12	1456
삼승면	일치수	31	78	18	22	45	22	18	9	5		248
	총수	103	239	90	115	325	351	483	389	134	18	2247
수한면	일치수	21	44	14	18	32	24	31	26	18	2	230
	총수	74	108	72	96	172	242	334	329	136	21	1584
외속리면	일치수	20	56	15	21	43	40	45	16	8	2	266
	총수	55	110	54	76	130	162	220	174	65	9	1055
탄부면	일치수	29	68	48	37	88	86	97	74	28	4	559
	총수	76	144	98	85	191	279	393	352	139	17	1774
회남면	일치수	10	8	5	4	6	5	2	4	1		45
	총수	26	30	13	29	53	82	111	98	38	4	484
회북면	일치수	31	41	16	14	29	20	23	6	3		183
	총수	71	108	60	78	239	286	354	360	130	24	1711
총합계	일치수	540	884	302	378	564	404	354	210	82	17	3735
	총수	2289	3147	1514	2617	4111	4039	4736	4034	1404	180	28115

<표 5-29> 방법2를 이용한 연령별 매칭결과

읍면동	10세 미만	10 - 19	20 - 29	30 - 39	40 - 49	50 - 59	60 - 69	70 - 79	80세 이상	총합계
내북면	30.8	38.9	17.9	10.8	12.6	13.1	5.3	4.8	5.9	11.7
내속리면	44.8	40.1	24.7	29.2	16.8	10.5	7.5	4.2	1.4	17.6
마로면	35.4	39.8	38.6	21.8	18.1	11.0	6.2	3.8	2.3	14.8
보은읍	16.7	18.7	13.3	9.5	8.0	5.9	3.2	2.3	2.1	9.6
산외면	37.2	42.3	23.5	23.1	17.9	7.9	6.5	3.4	1.2	13.6
삼승면	30.1	32.6	20.0	19.1	13.8	6.3	3.7	2.3	3.7	11.0
수한면	28.4	40.7	19.4	18.8	18.6	9.9	9.3	7.9	13.2	14.5
외속리면	36.4	50.9	27.8	27.6	33.1	24.7	20.5	9.2	12.3	25.2
탄부면	38.2	47.2	49.0	43.5	46.1	30.8	24.7	21.0	20.1	31.5
회남면	38.5	26.7	38.5	13.8	11.3	6.1	1.8	4.1	2.6	9.3
회북면	43.7	38.0	26.7	17.9	12.1	7.0	6.5	1.7	2.3	10.7
총합계	23.6	28.1	19.9	14.4	13.7	10.0	7.5	5.2	5.8	13.3
이내성(2006)	59.3	56.0	45.8	39.5	37.3	30.3	21.4	18.1	14.0	34.4



(그림 5-46) 방법2를 이용한 매칭의 일치율 비교

나. 부산광역시 해운대구 우동

(1) 방법1

방법1을 이용한 부산광역시 해운대구의 우동 자료의 매칭 결과는 다음과 같다. 시범조사 자료에서 조사받은 인구는 51,512명이다. 이 가운데 성별 없는 개체 포함되어 이들을 빼면 <표 5-30>과 같이 51,506명이 된다. 한편, 주민등록자료는 50,000명이다. 시범조사 자료의 51,512명 중에서 주민등록자료와 생년월일, 성별, 주소가 같은 사람이 33,584명이고 이는 시범조사를 받은 사람의 65.2%에 해당한다.

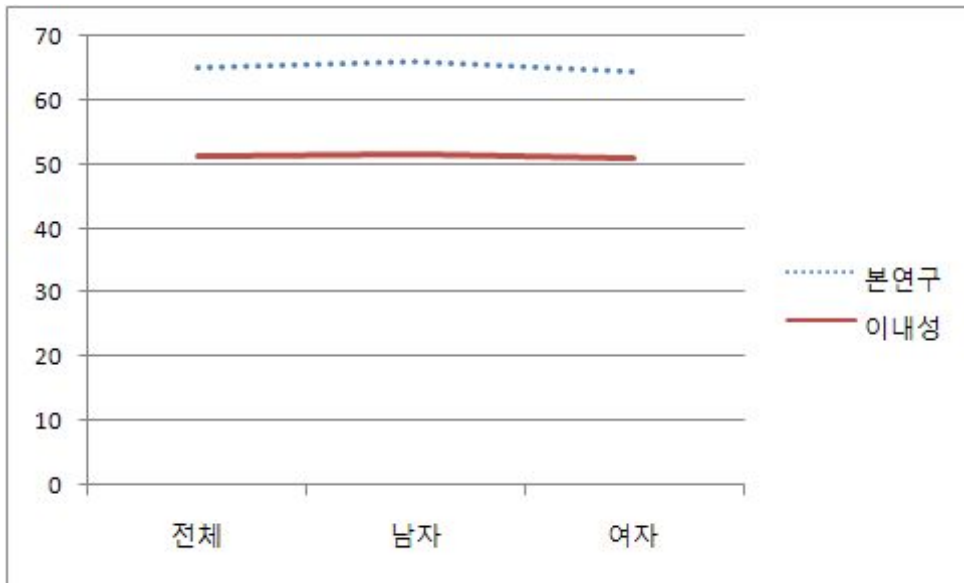
이내성(2006)과의 차이에 관하여는 매칭방법의 차이 때문인지, 주민등록자료의 가공 문제인지 판단을 할 수 없다. 특히, 본 연구에서 사용한 우동의 주민등록 자료의 크기는 딱 떨어지는 50,000인데 이것은 자료파일을 가공한 것이라는 의심이 든다. 그러나 전체, 남자, 여자 모두 약 15%의 차이가 나타난 것으로 보면 방법의 차이로 추측이 되기도 한다. 시범조사 자료를 기준으로 한 것이기 때문에 이내성(2006)의 시범조사 자료의 크기가 약 7,000명 정도 많았던 것을 고려하면 매칭률은 57.4%(33,584 / 58441)로 차이가 6-7%난 것으로 볼 수 있다.

<표 5-30> 방법1을 이용한 매칭결과

	주민등록 인구			시범조사 인구			일치 결과			일치율(%)		
	전체	남자	여자	전체	남자	여자	전체	남자	여자	전체	남자	여자
본 연구	50,000	24,579	25,421	51,506	24,878	26,628	33,584	16,410	17,174	65.2	66.0	64.5
이내성(2006)	54,225	26,606	27,619	58,441	28,691	29,750	29,942	14,803	15,139	51.2	51.6	50.9



(그림 5-47) 방법1을 이용한 매칭의 일치수 비교



(그림 5-48) 방법1을 매칭의 일치율 비교

연령별 매칭을 보면 다음의 표와 같으며 시범조사 자료와 행정자료에서 기타에 해당하는 부분은 생년월일을 입력하지 않았거나 입력은 했으나 연도 구분이 불가능(행정자료의 경우 엑셀 파일로 자료가 존재하므로 2003년도에 태어난 경우 03으로 입력되지만 엑셀의 특성상 0이 생략되고 3만 입력되어

있어서 2003년생인지 1903년생인지 구분이 불가능)한 자료 등에 해당한다.

이내성(2006)은 해운대구 전체에 대하여 표로 수록하였기 때문에 우동을 비교할 수는 없었는데 본 연구결과에서 우동만을 보면 30대가 가장 높은 일치율을 보였고, 나머지 70대 미만은 거의 고르게 나타난 반면, 이내성(2006)의 해운대구 전체에서는 40대까지와 50대 이상으로 차이가 있는 것으로 나타났다.

<표 5-31> 방법1을 이용한 매칭결과

연령대	주민등록 인구	시범조사 인구	일치 결과	일치율 (%)	이내성(2006)	
					일치 결과	일치율 (%)
10세 미만	2,492	4,801	2,524	52.6	19,074	44.4
10 - 19	5,953	7,007	4,641	66.2	26,510	44.6
20 - 29	6,716	6,760	4,816	71.2	23,572	43.0
30 - 39	6,885	8,140	5,416	66.5	28,345	43.1
40 - 49	7,926	9,170	6,413	69.9	31,624	42.2
50 - 59	7,643	7,631	5,106	66.9	15,495	32.0
60 - 69	5,371	4,828	3,183	65.9	6,788	24.0
70 - 79	2,939	2,374	1,205	50.8	1,989	14.1
80세이상	1,840	669	280	42.5	207	5.3
기타	2,235	132	-	0.0		
총인구	50,000	51,512	33,584	65.2	15,3604	39.1

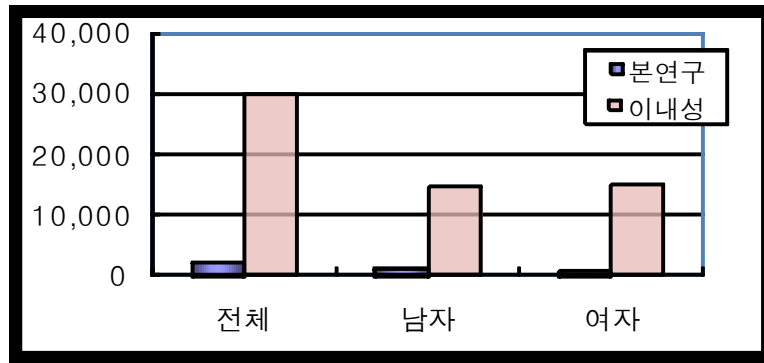
(2) 방법2

방법2를 이용한 부산광역시 해운대구의 우동 자료의 매칭 결과는 다음과 같다. 시범조사 자료에서의 인구는 51,512명이고 주민등록자료는 50,000명이다. 시범조사 자료의 51,512명 중에서 주민등록자료와 생년월일, 성별, 주소가 같은 사람이 2,058명이고 4%에 해당한다. 이 값들은 매칭률의 하한으로 볼 수 있다. 왜냐하면 한 주소에 두 사람이상이 매칭되었을 때만 매칭된 것으로 보는 것으로 정확성을 가장 중시하기 때문이다. 실제 매칭률은 4%보다는 높다고 할 수 있다.

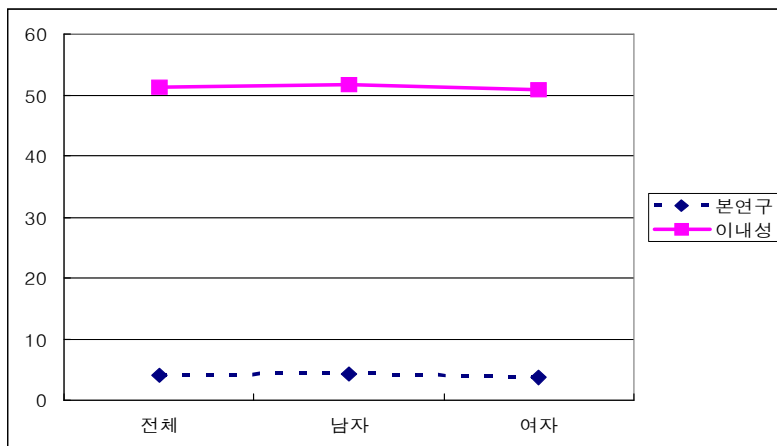
이내성(2006)과의 차이에 관하여는 매칭방법의 차이 때문인지, 주민등록자료의 가공 문제인지 판단을 할 수 없다. 그러나 전체, 남자, 여자 모두 약 46%의 차이가 나타난 것으로 보면 방법의 차이로 추측이 된다.

<표 5-32> 방법2를 이용한 부산광역시 해운대구 우동의 매칭결과

	주민등록인구			시범조사인구			일치 결과			일치율(%)		
	전체	남자	여자	전체	남자	여자	전체	남자	여자	전체	남자	여자
본 연구	50,000	24,579	25,421	51,506	24,878	26,628	2,058	1,073	985	4	4.3	3.7
이내성(2006)	54,225	26,606	27,619	58,441	28,691	29,750	29,942	14,803	15,139	51.2	51.6	50.9



(그림 5-49) 방법2를 이용한 매칭의 일치수 비교



(그림 5-50) 방법1을 이용한 매칭의 일치율 비교

연령별 매칭을 보면 다음의 표와 같으며 시범조사 자료와 행정자료에서 기타에 해당하는 부분은 생년월일을 입력하지 않았거나 입력을 했으나 년도 구분이 불가능한 자료 등에 해당한다.

이내성(2006)은 해운대구 전체에 대하여 표로 수록하였기 때문에 우동을 비교할 수는 없었는데 본 연구결과에서 우동만을 보면 30대가 가장 높은 일치율을 보였고, 나머지 70대 미만은 거의 고르게 나타난 반면, 이내성(2006)의 해운대구 전체에서는 40대까지와 50대 이상으로 차이가 있는 것으로 나타났다.

<표 5-33> 방법2를 이용한 매칭결과

연령대	주민등록 인구	시범조사 인구	일치 결과	일치율 (%)	이내성(2006)	
					일치 결과	일치율 (%)
10세 미만	2,492	4,801	167	3.5	19,074	44.4
10 - 19	5,953	7,007	479	6.8	26,510	44.6
20 - 29	6,716	6,760	311	4.6	23,572	43.0
30 - 39	6,885	8,140	278	3.4	28,345	43.1
40 - 49	7,926	9,170	290	3.2	31,624	42.2
50 - 59	7,643	7,631	213	2.8	15,495	32.0
60 - 69	5,371	4,828	186	3.9	6,788	24.0
70 - 79	2,939	2,374	94	4.0	1,989	14.1
80세이상	1,840	669	40	6	207	5.3
기타	2,235	132	-	0.0		
총인구	50,000	51,512	2,058	4	15,3604	39.1

(3) 번지 내 세대수에 의한 검토

이 절에서는 동일번지 내의 세대수가 하나인 경우와 둘인 경우에 어떤 차이가 나타나는지를 검토하여 보았다. 부산광역시 해운대구 우동에는 현재 본 연구에서 사용하고 있는 주민등록자료를 기준으로 하면 50,000명의 주민이 11,752 세대(주민등록자료에서 세대주와의 관계가 본인 사람의 수 - 세대주

수)를 이루며, 16,204개의 번지에 거주하고 있다.

이중에서 한 번지 내 한 세대만 살고 있는 번지수는 11,258이고 그들의 인구는 31,839명이다. 반면 한 번지 내 두 세대가 살고 있는 번지수는 1,415이며, 그들의 총인구는 6,454명이다. 이들을 각각의 자료를 시범조사 자료에 매칭시킨 결과, 번지 내 한 세대만 살고 있는 주민이 매칭된 시범조사 자료의 가구는 8,394가구에 18,490명이고, 번지 내 두 세대가 함께 살고 있는 주민들을 매칭했을 경우는 3,344가구에 4,349명으로서 이를 표로 정리하면 <표 5-34>와 같고, 연령대와 연결하여 보면 <표 5-35>, <표 5-36>과 같다.

<표 5-34> 부산광역시 해운대구 전체의 매칭결과

	세대수(번지)	주민등록 인구	일치 결과	주민등록자료 기준의 일치율
1번지 1세대	11,258세대(11,258번지)	31,839명	8,394가구(18,490명)	26%
1번지 2세대	12,830세대(1,415번지)	6,454명	3,344가구(4,349명)	67%
우동 전체	11,752세대(16,204번지)	50,000명		

<표 5-34>에 나타난 것과 같이 한 번지에 한 세대만 있는 경우가 한 번지에 두 세대가 있는 경우보다 일치율이 크게 떨어지는 현상은 특이하다. 여기서 세대는 주민등록자료 기준, 가구는 시범조사 기준이기 때문에 직접적인 비교는 불가능하지만 향후 연구 과제를 제시하고 있다.

이러한 현상으로부터 의미 있는 요인을 찾아낼 수 있다면 주민등록자료의 실태파악, 또는 시범조사의 사후조사 등에 상당히 유용한 정보를 추출할 수 있다고 판단된다. 그러나 본 연구에서는 주민등록자료 자체의 정확성이 거의 보장되지 않는 상황이기 때문에 더 이상의 연구는 할 수가 없었다.

4. 소결 및 시사점

본 연구를 수행하는 과정에서 가장 크게 느낀 점은 시범조사의 시범조사 기획단계에서 세운 조사목적에 본 연구와 같은 매칭개념이 포함되지 않았다

고 추측된다. 조사 자료나 주민등록자료가 개인정보를 많이 포함하고 있고, 또 다른 부처의 자료이기 때문에 많은 어려움이 있겠지만, 만약 시범조사에서만 이것들을 특별 관리하여 본 연구와 같은 취지를 갖는 작업을 한다고 하면 관련 부처의 담당부서를 설득시킬 수 있으리라 생각된다. 따라서 시범조사의 목적중 매칭을 통한 주민등록자료와 비교를 우선순위가 높은 항목으로 넣는 것을 제안한다.

(1)항을 좀 더 구체적으로 기술하면 주민등록자료에 있는 성명, 주소가 시범조사 자료에 나타나지 않기 때문에(시범조사 자료에는 주소가 읍면동까지만 나와 있어서 도움이 되지 않는다) 결국 각 읍면동 내에서 성별과 생년월일을 사용하여 매칭을 시키게 된다. 그러나 시험자료에서의 생년월일은 양력, 음력, 생활에서 사용되는 생일 등으로 낮은 정확도를 갖고 있고, 주민등록자료에서의 생년월일은 양력, 음력인지 구별이 되지 않기도 하고, 시범조사에서 응답한 것과 다를 가능성도 높기 때문에 생년월일을 매칭에서 사용한다는 것은 사실은 적절하지가 않다. 따라서 본 연구나 이내성(2006) 연구에서나 산출된 일치율은 과장된 값이라고 볼 수도 있고 또한 반대로 볼 수도 있다. 불일치 정도를 조사하는 것 역시 정답이 없기 때문에 본 연구에서는 알아낼 수가 없다.

본 연구진행 중에 연구에 필요한 자료의 수집, 전달과정에서 느낀 것은 시범조사에 대한 중요도가 실제 조사업무에 대하여 주로 집중되어 있었다고 추측된다. 추측의 이유는 자료를 요청, 전달받는 과정에서 조사 자료에 관한 메타자료의 전달이 쉽지 않았고, 행정자료(타 부처 자료이기 때문에 여러 가지 제약이 있었다고 판단됨)의 관리가 정확하게 이루어지지 않고 있었다는 점이다 또한 연구자의 입장에서는 시범조사이지만 주소DB 등이 있을 수 있다고 추정되었는데 구할 수가 없었던 것 역시 시범조사의 목적이 너무 한쪽으로 치우쳐 있었다고 판단하게 된다.

행정자료 공유 등과 같은 시범조사에 대한 여러 가지 대안에 대한 검토 요구와 나아가 구체적 변경에 대한 요구가 증대하고 있는 상황에서 현재 시범조사의 조사 자료와 주민등록자료의 섬세한 매칭작업의 필요성은 매칭작업이 잘 이루어진다면 얻을 수 있는 다음과 같은 결과들에서 중요하다고 본

다.

매칭 된 사람들과 매칭이 안되는 사람들에 대한 분석(주민등록자료 기준, 시범조사 기준 양쪽에서 모두 수행)은 주민등록자료만 있을 때 그들에 대한 실제 주거확률 등을 추정할 수 있게 한다.

세대의 특성과 매칭 된 사람의 특성(매칭이 안 된 사람의 특성)간의 분석을 통하여 세대로 작성되어 있는 주민등록자료를 가구개념으로 추정해 나갈 수 있는 통계적 방법론 개발이 가능하다.

위와 같은 작업이 상이한 지역에서 이루어진다면 지역의 특성에 따른 추정치들의 변동을 분석하게 된다. 따라서 성명과 주소가 기록된 시범조사자료와 해당 지역의 동일 시점의 주민등록자료가 주어진다면 주민등록자료가 인구조사의 대안이 될 수 있는지 없는지의 모호한 논의를 본격화 할 수 있게 되리라 판단된다.

제6장. 인구주택총조사에서 건축물대장의 활용 가능성

제1절. 서론

이 장에서는 앞서 살펴본 (그림 4-5)에서 다룬 건축물대장으로 인구주택총조사의 조사항목을 대체할 수 있는지를 살펴보고자 한다. 이명진외(2006)는 건축물대장의 정보를 인구주택총조사에 활용할 수 있는지 알아보기 위해 그의 기본이 되는 인구주택총조사의 거처정보와 건축물대장을 각 건물 수준에서 매칭해 보았다. 실제 인구주택총조사 자료는 사용할 수 없었기 때문에 2004년에 실시한 인구주택총조사 5차 시범조사에서 획득한 가구명부와 거처정보과 해당 지역의 건축물대장을 이용하였다. 매칭에서 key는 지번(주소)을 사용하였으며, 수작업을 통해 상응하는 주소에 건축물을 찾아 매칭결과를 1대 1, 1대 다, 다대 다 등 여러 유형으로 분류하였다. 그런 다음 유형별로 건축물대장을 활용하는 방법을 직접 활용, 건축물 동수 구별 후 활용, 공동주택 확인 후 활용, 부속건물인 경우 활용, 대조자료로 활용, 활용 불가능으로 나누어 제안하였다.

이들의 연구는 건축물대장을 총조사에 활용할 때 어떻게 하면 활용성을 높일 수 있을 것인가에 초점을 둔 것이다. 그렇지만 제시한 유형이 두 자료가 상응하는(또는 상응하지 않는) 결과를 알고 난 후에 것이어서 건축물대장만을 갖고 총조사 자료를 대체하는 데는 매우 큰 한계가 있다. 다시 말해, 이 연구는 두 자료 사이에 일치와 불일치의 유형 분류로서 의미가 크지만 건축물대장으로 총조사를 대체하는 데 직접적으로는 큰 도움이 되지 않는다.

건축물대장의 정보가 총조사에 직접 활용 될 수 있으려면 건축물대장이 실제 건축물, 특히 주거용 건축물에 대한 정보를 실제와 다르지 않게 갖고 있어야 한다. 그러한 정보는 위치 정보에서부터 대지면적이나 용적률 등과 같은 건축물에 관련된 속성 정보를 포함하고 있다. 다시 말해, 건축물대장을 총조사에 활용하기 위해서는 건축물대장에 대한 연구가 있어야 한다.

이런 측면에서, 여기서는 먼저 건축물대장 대해 살펴보고, 이와 더불어 건

축물과 관련 있는 기타의 행정자료도 간략히 살펴본다. 다음은 건축물대장의 위치정보와 속성자료를 검토해 본다. 이때는 본 연구진의 한 연구원이 참여한 건축물대장과 관련 있는 기존 연구(강영옥·이주일, 2005)에서 이용한 자료와 분석결과들을 이용하고자 한다.

제2절. 건축물 관련 행정자료

1. 건축물대장

가. 건축물대장의 개요³²⁾

건축물대장은 적법하게 건축된 건축물에 대하여 건축과/주택과에서 사용승인한 후 지적과(시민봉사실)에서 작성 및 발급하는 공부이다. 대장은 소유권의 구분에 따라 일반건축물대장과 집합건축물³³⁾대장으로 나누어 작성된다. 주택과 같이 건축물의 부분 부분이 각기 구별된 소유권으로 나누어지지 않는 건축물을 일반건축물이라고 하며, 아파트나 상가와 같이 건축물의 부분이 각기 구별된 소유권으로 구성된 건축물을 집합건축물이라고 한다.

- 일반건축물대장은 대지, 건축, 연면적, 건폐율과 같은 기본사항과 층별 건축물현황, 소유자현황 등을 기재
- 집합건축물대장은 건축물의 기본사항과 건축물현황을 기록하는 표 제부와 소유자 현황을 기재하는 전유부로 구분
- 일반건축물이든 집합건축물이든 2이상의 건축물이 하나의 대지 안에 있는 경우에는 건축물대장 총괄표제부를 작성

건축물대장의 작성 대장별 상세 내역을 살펴보면 <표 6-1>, (그림 6-1), (그림 6-2) 등과 같다.

32) 윤혁경(2004), 한국전산원(1996), 행정자치부(1998)을 근거로 작성

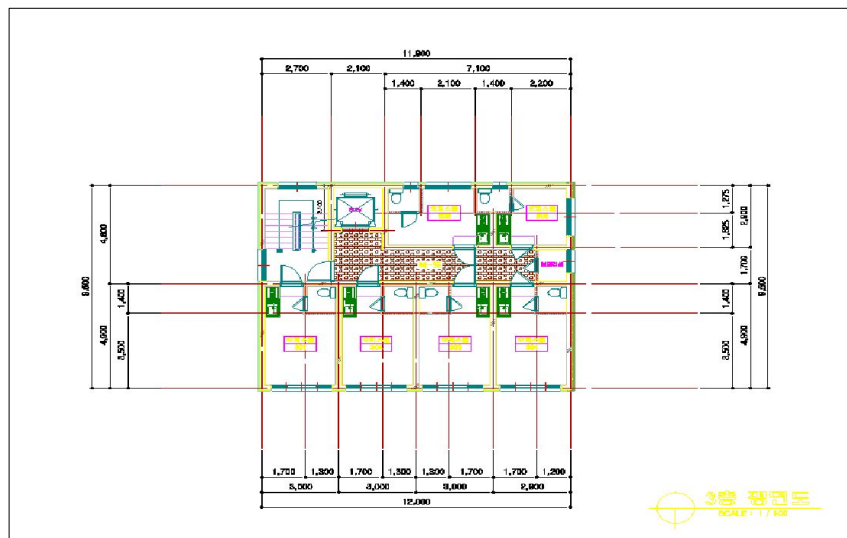
33) 집합건축물의 소유 및 관리에 관한 법률의 적용을 받는 건축물: 1동의 건물 중 구조상 구분된 수개의 부분이 독립한 건물로서 사용될 수 있을 때에는 그 각 부분은 이 법이 정하는 바에 따라 각각 소유권의 목적으로 할 수 있다(집합건축물의 소유 및 관리에 관한 법률 제1조의 건물의 구분소유).

<표 6-1> 건축물대장의 종류

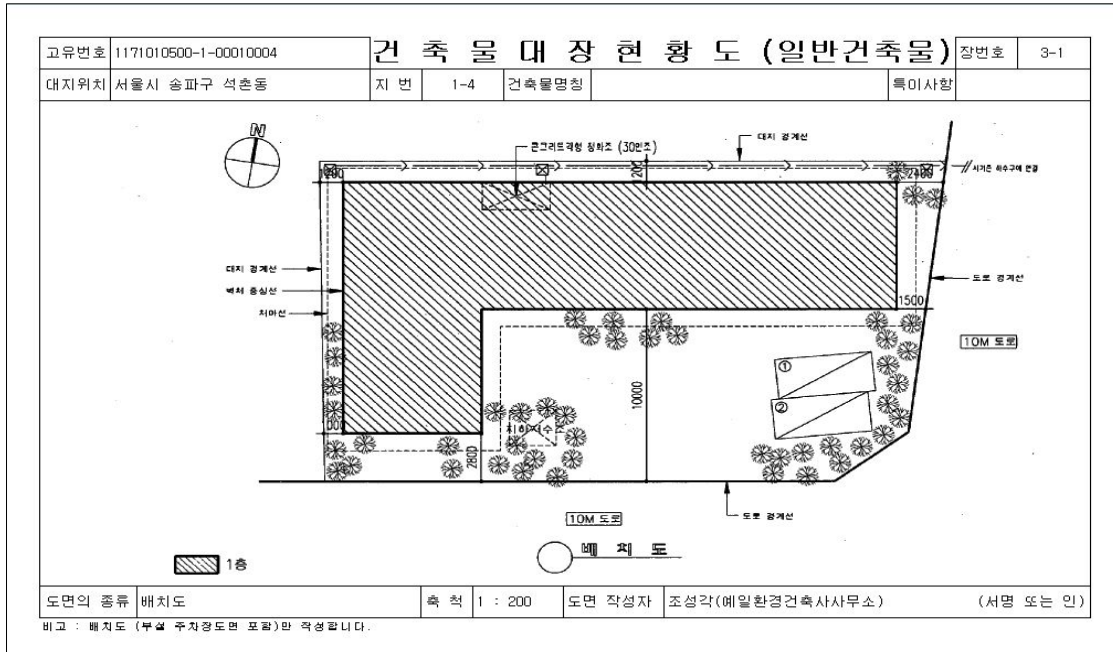
구분	개요	주요내용	첨부도면
일반 건축물 대장	<ul style="list-style-type: none"> · 소유권이 호별로 분리되지 않은 일반건축물을 관리하는 대장 · 집합 건물의 소유 및 관리에 관한 법률의 적용을 받는 건축물(이하 "집합건축물"이라 한다)외의 건축물 및 대지에 관한 현황을 기재한 건축물대장 	<ul style="list-style-type: none"> · 기본사항 · 건축물현황 · 소유지현황 	<ul style="list-style-type: none"> · 평면도, 배치도(총괄표제부에 배치도 존재시 생략)
집합 건축물 대장 (표제부)	<ul style="list-style-type: none"> · 호별로 소유권이 분리된 집합건축물의 현황을 관리하는 대장 · 집합건축물에 해당하는 건축물 및 대지에 관한 현황을 기재한 건축물대장 	<ul style="list-style-type: none"> · 기본사항 · 건축물현황 	<ul style="list-style-type: none"> · 평면도, 배치도(총괄표제부에 배치도 존재시 생략)
집합 건축물 대장 (전유부분)	<ul style="list-style-type: none"> · 호별 소유 현황을 관리하는 부분 · 건축물의 전유부분 기재 서식 	<ul style="list-style-type: none"> · 기본사항 · 전유현황, 공유현황 · 소유지현황 	<ul style="list-style-type: none"> · 평면도
총괄 표제부	<ul style="list-style-type: none"> · 하나의 대지에 2이상의 건축물이 있는 경우(일반건축물, 집합건축물) · 하나의 대지에 2이상의 건축물(부속건축물을 제외)이 있는 경우 작성하는 서식 	<ul style="list-style-type: none"> · 기본사항 · 건축물동별현황 	<ul style="list-style-type: none"> · 배치도

출처 : 건축물대장의 기재 및 관리 등에 관한 규칙(2004.11.29 개정 기준)

(그림 6-1) 건축물대장 관련 도면 예시-평면도



(그림 6-2) 건축물대장 관련 도면 예시-배치도



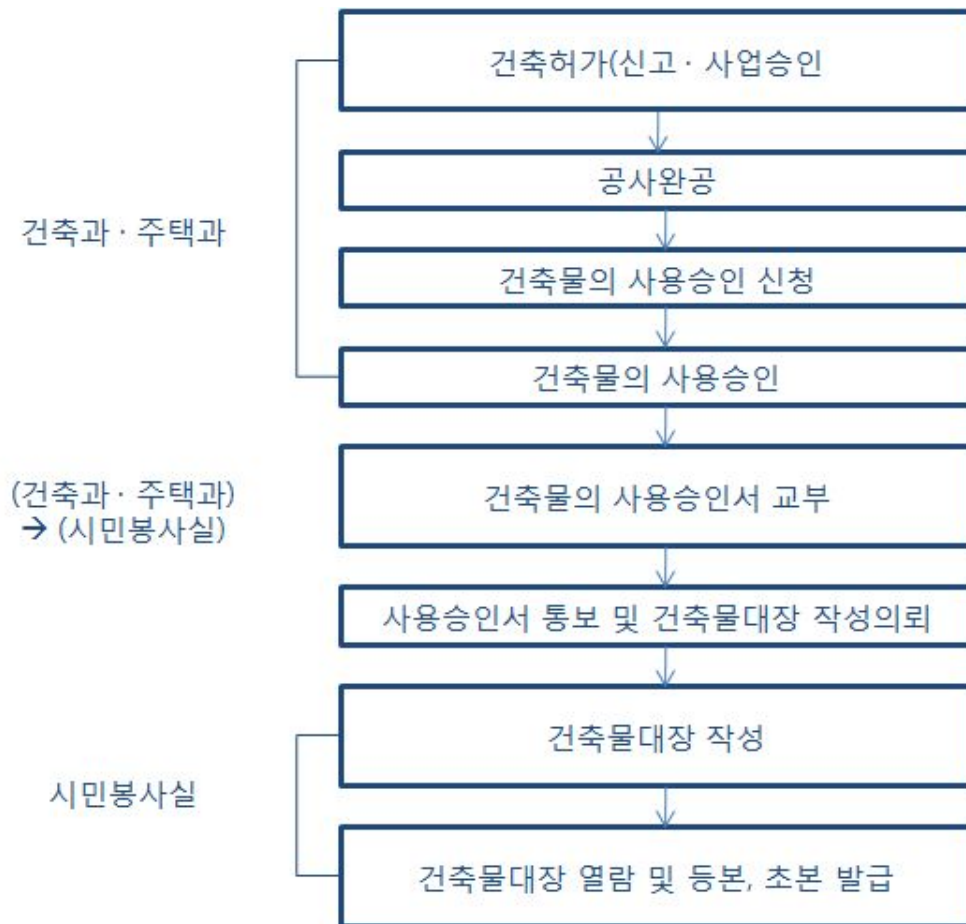
건축물대장은 다음과 같은 3가지 경우에 작성되며, 건축물대장의 작성절차는 (그림 6-3)과 같다.

- 건축물의 사용승인(신축, 증축, 개축, 재축, 이전, 대수선, 용도변경)후
- 직권에 의해성³⁴⁾
- 민원인에 의한 기재 신청 및 표시 변경에 의해 작성³⁵⁾

34) 직권에 의해 건축물대장이 작성되는 건축물 : 협의건축물(교육청, 군부대 등), 소유권이 국가 또는 자치단체에 있는 건축물

35) 건축물대장 기재대상

- 사용승인서를 교부한 경우(건축법 제 29조)
- 건축물의 공사를 완료한 후 그 건축물에 대하여 기재의 요청이 있는 경우(건축허가대상건축물, 신고대상건축물 제외)(건축법 제 29조)
- 가옥대장의 신규등록 및 변경등록의 신청이 있는 경우(건축법 시행령 제 25조)
- 적법 건축물 소유자가 건축물대장으로 이기신청을 하는 경우(건축법 시행령 제 25조)
- 기재내용 변경등의 필요가 있는 경우(건축법 시행령 제 25조)



(그림 6-3) 건축물대장 작성 절차
 출처 : 윤혁경(2004)

건축물을 건축하고자 할 때는 시장·군수·구청장의 허가(경우에 따라서는 신고)를 받도록 법에서 규정하고 있으며, 허가대상 건축물과 신고대상 건축물을 살펴보면 <표 6-2>와 같다. 건축물대장은 초기에 건축허가를 받은 건축물의 현황 위주로 관리하다가 건축물등록제도의 도입으로 건축물의 소유와 이용상태에 대한 관리로 변화하여 왔다. 건축물 등록제도가 정착되기 이전에 관련 법규가 여러 차례 바뀌었으며, 그 때마다 서식과 기재내용 등이 변하였다.

<표 6-2> 건축허가·신고 대상

건축허가대상		건축신고대상	
1. 국토의 계획 및 이용에 관한 법률에 의한 도시지역안의 건축 및 대수선		1. 바닥면적의 합계가 85㎡이내인 증축·개축·재축	
2. 제2종지구단위계획구역안에서의 건축 및 대수선		2. 연면적의 합계가 100㎡이하(순수한 단독주택은 330㎡이하)건축물의 건축	
3. 관리지역, 농림지역, 자연환경보전지역 등에서 건축하는 아래 건축 및 대수선 -200㎡ 이상이거나 3층 이상인 건축물(증축으로 당해 건축물이 200㎡이상이 되는 경우 포함)		3. 읍·면지역에서 건축하는 다음 건축물의 건축·대수선 -연면적 합계가 100㎡이하 주택(순수한 단독 주택은 330㎡이하) -연면적 200㎡이하 창고 -연면적 400㎡이하 축사·작물재배사	
4. 다음 시설에서 떨어지는 구역안의 건축 및 대수선		4. 건축물 높이 3m이하의 범위안에서 증축하는 건축물	
시설명	경계에서 떨어지는 거리	비고	
고속국도 철도 일반국도	100m 50m	눈에 보이지 아니하는 곳으로 시장·군수·구청장이 지정·공고한 구역 제외	
5. 지역의 균형적 발전 또는 지역계획 등을 위하여 시장·군수·구청장이 필요하다고 인정하여 지정·공고한 지역안의 건축 및 대수선		5. 표준설계도서에 의하여 건축하는 건축물로서 그 용도·규모가 주위환경, 미관상 지장이 없다고 인정하여 건축조례로 정하는 건축물	
6. 기반시설부담구역안의 건축 및 대수선		6. 공업지역, 산업단지 및 제2종지구단위계획구역(산업형에 한함)에서 건축하는 2층이하로서 연면적의 합계가 500㎡이하인 공장	

출처 : 윤희경(2004)

건축물대장의 주요 변천 내역을 살펴보면, 건축물 관리의 최초 형태는 1906년 토지 및 가옥증명제를 실시하면서 가옥세를 부과하고 징수하기 위해 활용한 가옥대장이었다. 가옥대장은 과세를 부과하기 위한 참고자료로서 법적 근거 없이 관리해오다가 1968년 가옥과세대장 운용요령(내무부 지세 1234-8471)이 시달되면서 과세를 부과하기 위한 참고자료로서 활용되기 시작하였다. 1979년 건축물관리대장 운영요령 시달 이후에는 과세를 위한 과세대장과 소유권 증명에 활용하기 위한 건축물관리대장으로 구분하여 건축물을 관리하였다. 이후 건축물의 이용실태와 건축정책기초자료로 활용하기 위하여 건축물관리대장을 건축물대장으로 개편하고 건축물대장의 서식, 기재내용, 절차 등을 규정하였다.(건축물대장의기재및관리등에관한규칙, 건설부령

제 507호, 1992. 6.1)

시기별로 건축물대장의 기재항목도 달라졌으며, 관련 업무에서 주로 활용하는 주요 항목의 기재내용을 비교하면 다음과 같다.

- 1979년 이전 대장 : 건축물면적, 구조, 용도, 준공검사년월일
- 1979년-1992년 대장 : 건축면적, 건축물명칭·번호, 건폐율, 구조, 구조 및 지붕, 부속건축물, 연면적, 용도, 용적률, 주소, 준공일자, 지역지구, 층수
- 1992년 이후 대장 : 건축면적, 건축물명칭, 건축물 주구조, 건폐율, 대지면적, 부속건축물, 연면적, 용도, 용적률, 용적률산정용연면적, 주용도

나. 건축물대장의 정보

건축물대장의 정보는 건축물의 현황을 파악할 수 있는 위치정보, 층수, 높이 등의 기본현황정보와 소유 정보로 구성되어 있다. 건축물대장별 주요 항목과 이의 전산화 현황을 살펴보면 <표 6-3>와 (그림 6-4)와 같다.

<표 6-3> 건축물 대장별 주요 항목

일반건축물대장	집합건축물대장 (표제부)	집합건축물대장 (전유부분)	총괄표제부
<ul style="list-style-type: none"> • 고유번호 • 대지위치 • 지번 • 명칭 및 번호 • 대지면적, 연면적, 지역, 지구, 구역, 건축면적, 용적률 산정용 연면적, 주구조, 주용도, 층수, 건폐율, 용적률, 높이, 지붕, 부속 건물 동수 면적, 층별 구조·용도·면적, 특이사항 • 건축주 성명 및 면허번호, 공사감리자 성명 및 면허번호, 공사시공자 성명 및 면허번호 • 주차장 대수와 면적(옥내/옥외, 자주식/기계식), 승강기대수, 오수정화시설 형식과 용량, 허가일자, 착공일자, 사용승인일자 • 층별 소유자 현황 : 소유자 성명·소유자 주소·소유권지분·소유자 변동일자와 원인 • 건축물현황도 	<ul style="list-style-type: none"> • 고유번호 • 대지위치 • 지번 • 명칭 및 번호 • 대지면적, 연면적, 지역, 지구, 구역, 건축면적, 용적률 산정용 연면적, 주구조, 주용도, 층수, 건폐율, 용적률, 높이, 지붕, 부속건축물 동수와 면적, 층별 구조·용도·면적, 특이사항 • 건축주 성명 및 면허번호, 설계자 성명 및 면허번호, 공사감리자 성명 및 면허번호, 공사시공자 성명 및 면허번호 • 주차장 대수와 면적(옥내/옥외, 자주식/기계식), 승강기대수, 오수정화시설 형식과 용량, 허가일자, 착공일자, 사용승인일자 • 변동사항 및 변동내역 • 건축물현황도 	<ul style="list-style-type: none"> • 고유번호 • 대지위치 • 지번 • 명칭 및 번호 • 호명칭 • 전유부분현황 : 층별 구조·용도·면적 • 공유부분현황 : 층별 구조·용도·면적 • 소유자현황 : 소유자 성명, 주민등록번호, 주소, 소유권지분, 동일자, 변동원인 • 건축물현황도 	<ul style="list-style-type: none"> • 고유번호 • 대지위치 • 지번 • 명칭 및 번호 • 대지면적, 연면적, 지역, 지구, 구역, 건축면적, 용적률 산정용 연면적, 건축물 수, 주용도, 층수, 건폐율, 용적률, 층고, 주용도, 층수, 부속건축물 동수와 면적 • 건축물현황 : 건축물 명칭, 건축물 주구조, 건축물 지붕, 층수, 용도, 연면적, 변동일자, 변동원인 • 주차장 대수와 면적(옥내/옥외, 자주식/기계식), 승강기대수, 오수정화시설 형식과 용량 • 건축물현황도

(그림 6-4) 건축물 대장의 전산화된 형태 예시

다. 건축물대장 정비 사업

(1) 건축물대장 전산화 사업(행정자치부)³⁶⁾

건축물 대장 전산화사업은 가옥대장, 건축물관리대장, 건축물대장 등 수기로 관리되어 오던 각종 건축물 등록 대장을 표준화된 서식에 맞추어 전산화하는 것을 목적으로 행자부의 지적과가 주축이 되어 사업을 진행하였다.

건축물대장 전산화사업은 IMF직후 공공근로사업의 일환으로 실행되면서 건축물대장 자체의 정비에 초점을 두기보다는 기존 건축물대장을 전산화하는 것에 주된 목적이 있었다. 그 결과, 기존의 건축물 관련 대장의 항목 정비와 실태에 대한 검토 없이 표준화된 새로운 서식에 맞추어 전산으로 입력하여 표준화된 건축물대장 서식의 새로운 항목에 기재되는 내용들이 누락된 채 입력되는 문제가 있었다. 표준화된 건축물대장 서식의 항목은 기존 가옥대장이나 건축물관리대장의 항목보다 3배 이상 증가하였다.

36) 한국전산원(1996), 내무부·한국전산원·벽산엔지니어링(1998)

행정자치부를 중심으로 이루어진 건축물대장 전산화 사업은 건축물대장자료의 전산화와 건축물대장관리시스템³⁷⁾을 개발하는 것을 주요 내용으로 하였다. 시스템으로 개발된 대장관련 업무는 건축물 변경접수/정리, 소유권변경정리, 건축물기록조회, 건축물통계와 정책정보, 위법건축물관리 등의 업무를 대상으로 하였다.

(2) 건축행정정보화 사업(건설교통부)

건교부의 건축행정자료화 사업은 건축 및 주택분야의 업무를 전산화하기 위하여 추진된 사업으로, 전산화 대상 업무는 건축사무소, 시·도 및 시군구의 건축 및 주택관련 부서를 포함한다. 업무 전산화 시스템은 건축행정자료시스템과 건축정책정보시스템으로 구분되며 주요 내용은 다음과 같다.

- 건축행정자료시스템(Architectural Administration Information System) : 건축 인허가부터 착공, 사용승인에 이르는 건축행정업무와 주택 조합설립사업계획승인에서 사용검사에 이르는 주택행정업무의 전과정을 전산화하고, 처리과정에서 축적된 정보를 이용하여 건축물대장을 자동생성하고 전자도면 검색 및 각종 통계 제공
- 건축정책정보시스템(Architectural Administration Decision Support System) : AIS에 의해 축적된 건축·주택·건축물대장 전산자료를 이용하여 정책수립에 필요한 각종 통계를 손쉽게 작성하고 건축관련업자 행정처분자료 등 단체관리 통합체계 구축

(3) 건축물대장 정비사업(건설교통부)

부동산 투기에 따라 급변하는 주택시장을 안정화하기 위하여 건축주택에 대한 인별·세대별 보유현황을 정확하게 파악할 필요성이 대두되면서, 건설교통부는 건물, 등기, 토지자료를 일원화하여 건축행정의 종합적인 서비스 기반을 마련하고자 건축물대장 정비사업을 추진하였다. 건축물대장정비사업

37) 내무부·한국전산원·벽산엔지니어링, 1998, 건축물대장관리시스템 개발보고서

의 주요 내용은 건축물대장의 정비, 별도로 관리되어 오는 건축물대장과 등기부등본 자료를 연계하기 위한 공동고유번호 관리 프로그램 개발, 건축물대장 정비지침을 작성하는 것으로 구성되어 있다.

2. 기타 건축물 관련 행정자료

가. 과세대장³⁸⁾

건축물에 대한 과세를 산정하기 위하여 세무과에서는 과세대장을 관리하여 왔다. 2005년도부터는 개별주택에 대한 과세를 현실화하기 위하여 주택특성조사를 통하여 개별주택가격을 산정하고 있는데, 상가나 업무시설은 2007년도부터 개별건물 특성조사를 실시할 예정이다. 세무과에서는 과세 산정을 위한 근거자료로 건축물에 대한 정보를 관리하고 있다.

재산세 과세대장은 토지 과세대장과 건축물 과세대장으로 구분된다. 이중 건축물 과세대장은 공익사업을 목적으로 비영리사업자가 사업에 직접 사용하는 부동산과 같은 비과세 건축물을 제외한 모든 건축물에 대하여 매년 5월 1일을 기준으로 재산가액(과세시가 표준액)을 부과하기 위해 서면 혹은 직접 조사하여 세무과에서 작성하는 자료이다. 건축물 과세대장의 항목은 대장 관리번호, 성명·주소 등을 포함하는 재산세 납세의무자 관리, 구조·용도·면적 등을 포함하는 재산세 물건관리로 이루어져 있다.

과세대장은 과세대상인 건축물에 대한 정보만을 관리하므로, 비과세 건축물 정보는 포함하고 있지 않으며, 건축물의 동이나 호별로 과세대장을 정리하는 것이 아니라 건축물별 납세자의 구분에 따라 과세대상 물건을 정리하는 특징이 있다. 즉 2층짜리 단독 주택이면서 소유자가 한명인 경우는 한 건으로 처리되지만, 호별로 소유자가 나누어지는 아파트 같은 경우는 호별로 관리되고 있다. 또한, 과세대장은 과세를 부과하기 위한 목적으로 건축물의 용도와 구조를 나누고 있어 건축물대장과는 용도·구조의 분류체계가 다르다.

38) 서울특별시(2005a)

건설교통부에서는 부동산 보유세를 강화하고, 각종 토지 및 건물에 대한 세금을 현실화하기 위해 오래 전부터 실시해 온 토지특성조사와 유사하게 주택에 대해서도 2005년 처음으로 개별주택가격을 조사하여 관련 자료들을 전산화하였다. 개별주택가격을 산정하기 위해 주택특성조사³⁹⁾를 실시하여 개별주택에 대한 특성을 조사하고, 이를 근거로 주택가격을 산정하였다. 조사 대상은 단독, 다세대, 다가구, 주상복합등 주거용 건물이다. 주택특성을 조사하기 위한 기반자료로 건축물대장이 활용되며, 층별 용도 확인 등을 위해 재산세과세대장이 부가적으로 활용된다. 참고로, 주택특성조사는 매년 1월 1일(공시기준일)을 기준으로 하고, 토지의 분할·합병 및 건물의 신축 등이 발생한 단독주택은 매년 7월 1일을 기준으로 한다.

나. 위법건축물대장⁴⁰⁾

건축법에 적합하지 아니한 건축물을 위법건축물이라 하며, 이는 유휴가 위법 건축물과 신발생무허가건축물, 기존무허가건축물로 나누어진다. 유휴가 위법 건축물과 신발생무허가건축물에 대하여는 건축법에서 정한 행정적 처분이나 위반사항에 대한 시정명령을 내리며, 미 이행된 위법건축물에 대하여는 고발·단전단수와 이행강제금을 부과한 후에 건축물대장에 ‘위반건축물’로 표기하도록 되어 있다.

건축과의 점검과 단속을 통해 적발된 위법건축물은 (유휴가) 위법건축물대장과 신발생무허가건축물대장, 기존무허가건축물대장에 각각 기재된다. 유휴가 위법 건축물의 점검은 기존 건축물 중 중·대형 규모의 건축물과 신축 건물 중 소형건축물 및 다가구·다세대 주택, 건축신고, 가설건축물로 구분하여 그 점검시기와 점검방법, 점검내용을 달리하고 있다.

신발생무허가건축물에 대한 점검은 건축법의 허가 또는 신고를 거치지 않고 이루어진 건축물 및 가설건축물과 개발제한구역 내 구역지정일 이후 건축한 건축물을 대상으로 한다. 신발생무허가건축물 점검활동은 단속시점에

39) 건설교통부(2004), 서울특별시(2005c)

40) 서울특별시(2005b)

따라 무허가건축물이 완성되기 이전 단계에서 단속하는 예방단속과 무허가 건축물이 완성된 단계에 단속하는 사후조치로 나누어지며, 순찰보고제· 주민신고제와 같이 현장 확인에 의해 대상을 적발하거나 항공사진 판독을 통해 적발된 건축물을 대상으로 한다.

기존무허가건축물은 다음과 같은 경우에 해당하는 건축물을 의미하며, 기존건축물관리대장을 별도로 작성하여 관리하고 있다.

- 1981.12.31 현재 무허가건물관리대장에 등재된 건물
- 1981년 제2차 촬영 항공사진에 수록되어 있는 무허가건물
- 재산세 납부 등으로 공부상 1981.12.31. 이전에 존립하였다는 확증 있는 무허가건물
- 1982.4.8. 이전에 사실상 건립된 연면적 85㎡ 이하의 주거용 건물로서 1982년 제1차 촬영 항공사진에 수록되어 있거나 또는 재산세납부등 공부상 1982.4.8. 이전에 건립하였다는 확증이 있는 무허가 건물

위법 건축물대장에서 관리하는 항목은 대장별로 차이가 있으며, 대장별 관리 내역은 다음과 같다.

<표 6-4> 위법 건축물 관리 대장별 항목

• 위법건축물대장 항목																		
위치	건축주	건물규모		용도	위법내용	조치내용 (행정처분)	시정여부											
		층수	연면적					건축주고발, 이행강제금부과, 시정지시등 내용 기재										
• 신발생무허가 건축물 점검 항목																		
일련번호	도엽 및 건축물 번호	적출사항																
		접수 일자	적출 유형	위치	건축	면적 (㎡)	용도	구조	형태 및 공정 (%)	지구 구분	토지 소유 구분							
• 기존무허가 건축물대장 항목																		
일련번호	향측도	동일련번호		70.6.20 대장등재	위치	건물주			건물			토지			지구별			
	도엽번호	건물번호	동번			건물번호	성명	주민등록번호	구조	연면적 (㎡)	층수	용도	건축시기	면적	소유별	지목	도시계획	개발지구여부

다. 등기부등본(건물분)

부동산등기부는 무형적인 권리관계를 등기부에 기재하여 공시하는 제도로써 사법부인 법원행정처 주관 하에 산하기관인 지법과 지원 및 등기소에서 부동산등기법령에 의하여 수행하고 있다.

부동산등기부등본은 건물등기부등본과 토지등기부등본 2가지가 있으며 각각 표제부, 압구, 을구의 3부분으로 구성되어 있다. 을구에 기재된 사항이 전혀 없거나 기재된 사항이 말소돼 현재 효력이 있는 부분이 전혀 없을 때에는 을구를 제외한 표제부 및 압구만으로 구성된 등기부등본을 발급한다. 등기부 등본의 구성은 다음과 같다.

- 표제부 : 부동산의 표시에 관한 사항을 기재하는 곳으로 소재와 지번 구조 종류 및 면적 등을 알 수 있고, 아파트 등과 같이 대지권이 있는 경우에는 그 대지권의 표시에 관한 사항이 기재되어 있음
- 압구 : 소유권에 관한 사항을 기재하는 곳으로 소유자, 성명, 주소, 주민등록번호와 등기원인 및 그 년 월일을 알 수 있고 압류 가처분 가등기 사항을 알 수 있음
- 을구 : 소유권 이외의 권리사항을 기재하는 곳으로 근저당 전세권 등의 설정내용과 그 년월일 채권 채무자 등의 사항을 알 수 있으며 소유권 이외의 권리사항이 없을 경우에는 을구는 없음

<표 6-5> 등기부등본 정보

구분		상세내역
표제부	토지의 표시	표시번호, 소재지번, 지목, 면적, 등기원인 및 기타사항
	1동의 건물표시	표시번호, 접수, 소재지번, 건물명칭, 건물내역(구조,용도, 면적등)
(전유부)	대지권	표시번호, 대지권종류, 대지권비율, 등기원인 및 기타사항
	전유부분의 건물	표시번호, 접수, 건물번호, 건물내역, 등기원인 및 기타사항
압구(소유권에 관한 사항)		순위번호, 등기목적, 접수, 등기원인, 권리자 및 기타사항
을구(소유권외 권리 관한 사항)		순위번호, 등기목적, 접수, 등기원인, 권리자 및 기타사항

부동산 등기 전산화는 94년부터 97년까지 업무 재설계 및 응용시스템 분석, 설계 및 개발 작업을 수행한 바 있으며 전국의 등기소에 대해 2002년 9월말까지 전산화를 완료하였다. 건설교통부에서는 토지대장-건축물대장-등기부등본의 연계 작업을 수행한바 있으며, 이 사업으로 토지대장을 기반으로 건축물대장과 등기부등본의 지번을 일치시켰으며, 등기부등본을 기반으로 건축물대장의 소유자를 일치시켰다. 등기부등본은 소유권 관련 사항을 확인하기 위하여 사용되는 공부로서, 건물의 용도·구조·면적과 같은 건물 특성을 건축물대장과 중복적으로 기록관리 되고 있다.

라. 수치지형도

건물에 대한 도형 현황을 볼 수 있는 자료로는 서울시에서 구축한 1:1000 수치지형도의 건물도형레이어가 있으며, 새주소사업에서는 새주소 부여를 위해 신규로 구축되는 건축물에 대한 도형관리를 포함하고 있다. 서울시의 1/1000 수치지형도는 1996-1997년 항공사진을 기반으로 1996-1998년에 일차 구축이 완료되었으며, 전체 2,102도엽으로 정위치 편집된 도면단위 CAD(dwg, dxf)파일과 구조화 편집까지 완료된 레이어단위 Shape(shp)파일의 두 가지가 있다.

제3절. 건축물대장 활용의 문제

1. 건축물 대장의 문제점

가. 주요 항목누락

건축물 대장 기재항목 가운데 주요 항목이 기재되어 있지 않은 경우가 있어 정보로서 활용에 문제가 된다. 구별로 차이가 있기는 하지만 동용적용, 동건폐율, 동용적용산정시 연면적, 동높이, 동착공일자 항목의 누락율이 높게

나타났다. 건축물대장은 여러 번에 걸쳐 대장 작성 양식이 변경된 바 있기 때문에 건축물대장 정비사업을 추진한 '92년 이후 자료만을 대상으로 주요 항목별 누락율을 정리한 내용은 <표 6-6>과 같다.

<표 6-6> 건축물대장의 주요 항목별 누락률-92년도이후 사용승인분

테이블	주요항목	종로구		강남구		서대문구	
		누락수	비율(%)	누락수	비율(%)	누락수	비율(%)
동별 개요	동주용도코드	90	1.6	33	0.3	44	0.5
	동기타용도	12	0.2	33	0.3	-	0.0
	동구조코드	31	0.6	10	0.1	29	0.4
	동구조상세	9	0.2	10	0.1	1	0.0
	동지상층수	56	1.0	25	0.3	18	0.2
	동지하층수	30	0.5	16	0.2	67	0.8
	동용적률	693	12.5	594	6.0	1,103	13.8
	동건폐율	691	12.4	584	5.9	1,101	13.8
	동연면적	2	0.0	16	0.2	1	0.0
	동용적률산정시 연면적	3,195	57.5	5,943	59.9	5,393	67.4
	동건축면적	514	9.2	573	5.8	811	10.1
	동대지면적	334	6.0	540	5.4	1,091	13.6
	동허가일자	357	6.4	415	4.2	172	2.1
	동높이	676	12.2	898	9.1	938	11.7
	동지붕코드	355	6.4	186	1.9	206	2.6
	동지붕명	57	1.0	187	1.9	27	0.3
	동착공일자	981	17.6	1,223	12.3	1,096	13.7
	전체레코드수		5,560		9,920		8,003

건축물대장의 층별 개요자료와 총괄표제부의 항목별 누락률은 <표 6-7>과 같다.

<표 6-7> 건축물대장의 주요 항목별 누락률-층별자료, 총괄표제부

테이블	주요 항목	종로구		강남구		서대문구	
		누락수	비율	누락수	비율	누락수	비율
층별 개요	층주용도코드	15,670	16.3	2,299	1.5	11,916	10.6
	층기타용도	1,357	1.4	804	0.5	3,105	2.8
	층면적	186	0.2	30	0.0	80	0.1
	층번호	357	0.4	73	0.0	102	0.1
	층구조코드	17,064	17.8	5,999	3.9	3,046	2.7
	층구조상세	810	0.8	6,047	4.0	880	0.8
전체 레코드수		96,118		152,006		112,786	
총괄 표제부	주용도코드	989	88.3	6	1.9	994	83.0
	기타용도	1,015	90.6	7	2.3	1,049	87.6
	사용승인일자	1,103	98.5	282	90.7	1,182	98.7
	용적률	1,018	90.9	244	78.5	1,051	87.7
	건폐율	1,019	91.0	242	77.8	1,051	87.7
	연면적	1,013	90.4	4	1.3	1,044	87.1
	용적률산정시연 면적	1,028	91.8	262	84.2	1,062	88.6
	건축면적	1,018	90.9	244	78.5	1,049	87.6
	대지면적	1,014	90.5	242	77.8	1,051	87.7
	허가일자	1,106	98.8	288	92.6	1,183	98.7
	착공일자	1,111	99.2	287	92.3	1,188	99.2
전체 레코드수		1,120		311		1,193	

참고: 층별자료와 총괄표제부는 사용승인년도를 파악하기 어려워 전체를 대상으로 함

나. 건축물 대장 테이블 간 불일치 문제

건축물대장의 총괄표제부와 동별개요자료, 동별개요자료와 층별개요자료의 연계가 미흡하며, 자료별로 연계되어 있다 하더라도 연계된 정보간 내용이 상이한 경우가 있다((그림 6-5)~(그림 6-8)). 동별개요자료와 층별자료를

	A	R	S	T	U	V	W	X	Y	Z	AA
8											
9											
10	총괄표제부										
11								허가일자	착공일자	사용승인일자	세대수
12											745
13	동별개요										
14	동구조상세	동지붕코드	동지붕명	동높이	동지상층수	동지하층수	동허가일자	동착공일자	동사용승인일자	동세대수	동
15	철근콘크리트	10	평슬라브	66	24	1		19980320	20011201	96	
16	철근콘크리트	10	평슬라브	67.5	24	1	19980228	19980320	20011201	48	
17	철근콘크리트	10	평슬라브	56.3	20	1	19980228	19980320	20011201	40	
18	철근콘크리트	10	평슬라브	66	24	1	19980228	19980320	20011201	96	
19	철근콘크리트	10	평슬라브	61.7	23	1	19980228	19980320	20011201	88	
20	철근콘크리트	10	평슬라브	67.5	24	1	19980228	19980320	20011201	48	
21	철근콘크리트	10	평슬라브	66	24	1	19980228	19980320	20011201	96	
22	철근콘크리트	10	평슬라브	67.5	24	1	19980228	19980320	20011201	48	
23	철근콘크리트	10	평슬라브	66	24	1	19980228	19980320	20011201	96	
24	철근콘크리트	10	평슬라브	56.3	20	1	19980228	19980320	20011201	76	
25	철근콘크리트	10	평슬라브	10.2	2	1	19980228	19980320	20011201	9	

(그림 6-8) 허가일자/착공일자/사용승인일자가 다른 경우

다. 건축물 대장 자료관리의 문제점⁴¹⁾

(1) 건축물대장 작성 연도별 서식 유형이 다양하여 항목별 기재율이 다름

건축물대장 표준서식을 1996년도에 제정하였으나, 내용상의 표준이 이루어지지 않았으며 그에 따른 문제점은 다음과 같다.

- 1996년 이전 작성된 건축물대장의 문제점 : 건축물대장이 가옥대장, 과세대장, 건축물관리대장 등으로 변천되면서, 대장의 서식 개정이 있을 때마다 관리항목이 늘어나 기존에 있던 건축물 대장은 새로운 서식에 적용이 안되고 있어 누락항목이 많이 발생하고 내용적 보완이 안되는 문제가 발생함
- 1996년 표준화서식의 자료 기재 시 문제점 : 1996년 표준화 서식이 작성

41) 솔리데오시스템즈(2004), 건설교통부·솔리데오(2005)과 건축물대장 전산자료정비관련 업무 담당자와 업체담당자의 인터뷰를 근거로 작성하였다.

되면서 구 건축물대장을 표준화서식에 맞추어 이기하도록 하였으나 새로운 서식에 늘어난 항목에 대해서는 대부분 누락된 상태로 존재하며 대사 및 정비를 하여도 근거자료가 미흡한 상황임. 따라서 많은 자료들이 누락되어 자료 활용이 저조하며 자료에 대한 신뢰성도 적음. 또한, 구대장을 표준화 서식으로 이기할 때, 법·제 도적인 뒷받침이 없어 구대장에 있는 항목을 있는 그대로 표준화 서식에 이기하여 구대장 내용상으로는 별다른 차이가 없음.

- 총괄표제부가 누락되는 경우 다수 존재 : 제도적으로 건축물대장총괄표제부를 작성하게 있으나 총괄표제부가 작성되어 관리되는 경우는 극히 일부이며, 총괄표제부 없이 일반건축물의 동별 현황에 동별로 나열하여 관리하는 대장이 많음. 이런 대장이 많이 존재 하는 공단지역, 학교지역, 군부대 지역이며, 총괄표제부 없이 건축물 여러 동을 하나의 대장에 기재하는 경우 통계자료 작성 시 문제가 발생함

(2) 건축물대장자료의 이중관리

건축물대장 카드 자료와 전산자료가 이중으로 관리되고 있어 업무담당자에게 부담이 되고 있으며, 건축물대장의 변동이 발생할 경우 전산대장은 반영이 되나 카드대장은 누락되는 경우가 있으며, 반대로 카드대장은 수정이 되나 전산대장은 반영이 안 되고 누락되는 경우 등 각 자료가 일치하지 않는 문제점이 있다.

(3) 멸실 후 대장말소가 안되는 경우 발생

건축물이 신축되는 경우, 착공신고 이전에 기존의 건축물에 대한 철거·멸실신고를 하고 기존 건축물대장을 말소하여야 하며, 그런 후에 신축 건물에 대한 사용승인서를 참조하여 건축물대장을 작성하여야 한다. 기존 건축물대장이 말소처리가 되지 않고 존재하여 신축처리가 안 되는 경우가 종종 발생한다.

(4) 직권이나 민원인에 의한 기재변경 신청 시 문제점

건축물대장은 사용승인 후 작성되거나, 직권에 의해 혹은 민원인에 의한 기재 신청 및 표시 변경에 의해 작성되는데 직권에 의하거나 민원인에 의한 기재신청 및 표시 변경되는 경우 다음과 같은 문제점이 발생한다.

- 직권에 의해 작성되는 경우 문제점
 - 협의 건축물 (교육청, 군부대), 소유권이 국가 또는 자치단체에 있는 건축물에 대한 신축 및 표시변경(증축, 재축등)으로 건축물의 변동이 발생하는 경우 직권에 의해 건축물대장에 변동사항이 바로 적용되는데, 잘못 작성되거나 누락항목이 발생할 가능성이 매우 높음. 또한 직권에 의해 건축물대장에 변동사항이 바로 기록되기 때문에 대사 및 수정 시 근거 자료도 불충분함.
 - 무허가건축물이 양성화되는 경우 시군구 담당자에 의해 실사 없이 또는 실사가 정확하게 이루어지지 않은 상태에서 건축물대장이 작성되면서 건축물대장 항목이 누락되는 경우가 많이 발생함.
 - 건축법에 의해서가 아니라 관련법에 의해서 건축물대장이 신규 작성되는 경우, 건축물대장의 관리항목 사항을 정확히 기재하지 않고 건축물대장을 작성하는 경우가 발생함
- 민원인에 의한 기재 신청 및 표시 변경에 의해 작성되는 경우 문제점
 - 건축물의 기재 신청 및 표시변경사항을 접수 받아서 건축물 대장을 정리하나 기재신청서 및 표시변경 신청서에 기재된 내용을 그대로 입력하여 누락 및 잘못 기재된 경우가 발생

(5) 건축물대장의 표준화 문제

건축물대장 전산화 사업에서는 기존 건축물대장의 TEXT를 있는 그대로 입력하였다. 건교부의 건축물대장 이식·변화사업을 통하여 용도, 구조, 지붕 항목의 텍스트정보를 표준화하여 코드로 구축하였으나 전체 내용을 코드화

하기에 어려움이 있었다. 표준화하여 코드화하기 어려운 내용의 예로는 물치(창고), 육육근·평 옥계(평스라브기와, 변소·측간·해우소(화장실), 연와(벽돌조)등과 같이 현재 사용하고 있지 않은 용어이거나 축약어로 작성되어 의미를 파악하기 어려운 경우가 있었다.

라. 사용승인서를 기반으로 한 건축물 대장 정비 시 문제점

건축물대장 정비사업은 건축물에 대해 작성된 사용승인서를 기반으로 건축물대장을 정비하고 있다. 사용승인서를 기반으로 건축물대장을 정비하는 경우 사용승인서 자체가 가지고 있는 오류, 그리고 사용승인서가 작성되기 시작한 92년 이후의 건축물에 대해서는 정비가 가능하나, 92년도 이전에 작성된 건축물대장인 경우에는 누락된 항목의 보완이나 내용을 변경하는데 어려움이 있다.

- 사용승인서를 기초로 건축물대장을 정비할 때 문제점
 - 사용승인서를 참조하여 건축물대장이 신규로 작성되거나 건축물대장이 변동되는 경우 사용승인서의 서류는 정확한 반면 전산으로 처리되는 사용승인의 경우 사용승인서의 내용과 전산상의 사용승인의내용이 상이하여 건축물대장이 잘못 작성되는 경우 발생
 - 사용승인서를 참조하여 건축물대장의 전산자료를 정비하는 경우 사용승인서의 내용이 건축물대장의 서식에 적용하기 부적합하거나 누락되는 경우가 발생하여 사용승인서의 내용과 건축물대장의 내용이불일치하는 자료 발생
 - 사용승인과 함께 건축물대장이 작성되는데 사용승인 당시에는 일반건축물이었으나 그 이후 상가 및 오피스텔이 집합건축물로 전환되는 경우, 사용승인서를 참조하여 건축물대장을 대사·수정할 때 어려움이 발생
- 92년 이후 사용승인서 대조 건축물대장 정비 시 문제점⁴²⁾
 - 사용승인서의 주요 항목은 대지위치, 대지면적, 건축물명칭, 주용도,건축면적, 건폐율, 연면적, 용적률로 건축물대장의 누락 항목을 보완 할 수

42) 건설교통부. 2005. 「건축물대장 정비 및 검수 지침」.

있는 항목이 일부로 한정되어 있어 92년도 이후 건축물대장이라 할지라도 완벽하게 보완하는 데에는 어려움이 있음.

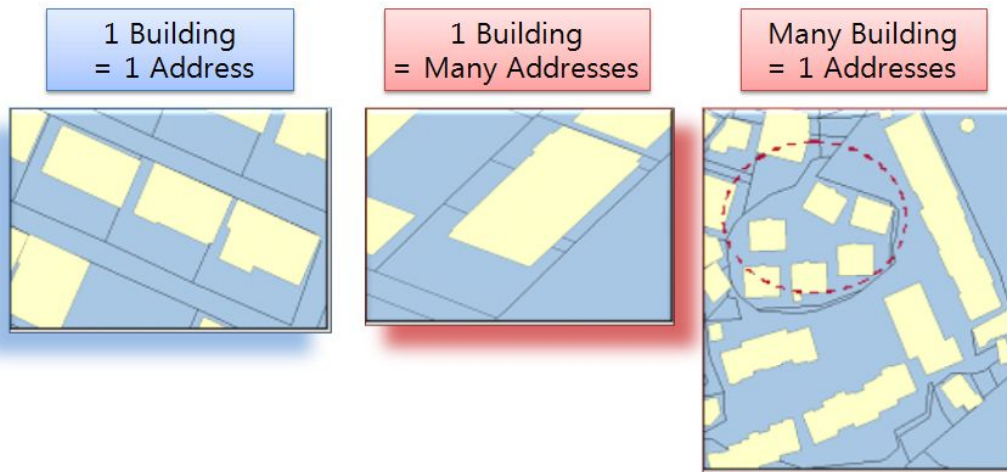
- 사용승인서 1건에 대해 다수의 건축물대장이 존재하거나, 1건의 건축물대장에 대하여 여러 개의 사용승인서가 존재하여 최종 사용승인서를 확인하기 어려운 경우, 사용승인서는 존재하나 건축물대장이 존재하지 않는 등 사용승인 서류와 건축물 대장의 비교가 어려운 경우가 있음
- 사용승인서류를 참조하여 건축물대장을 정리할 때 사용승인서의 내용과 건축물대장의 내용이 불일치하거나 건축물대장의 서식에 적용하기에 부적합하거나 누락되는 자료가 발생함.
- 건축물대장 작성 후 상가 및 오피스텔과 같은 건축물은 집합건축물로 전환되는 경우가 있는데 일반건축물로 신청한 사용승인서만을 참조하여 집합건축물로 변경된 건축물대장을 정비할 때 어려움이 있음.

2. 건축물대장과 관련 자료 연계의 문제

가. 건물관련 행정자료 연계키에 관한 개요

건축물의 위치정보는 건축물의 고유번호인 필지번호(지번주소)가 갖고 있다. 이 위치정보를 확인하는 것은 수치지형도의 건물도형을 통해서 할 수 있다. 이들 자료 사이의 연계는 건물도형도 편집지적도의 필지번호와 건물관련 행정자료의 고유번호인 필지번호를 활용한다. 이때 건물과 필지번호가 서로 1 대 1로 대응하지 않기 때문에 연계되지 않는 경우가 발생한다.

아래 (그림 6-9)는 필지번호와 건물의 관계 유형을 제시한 것이다. 첫 번째 유형은 건물과 필지가 1 대 1로 연계되어 문제가 없다. 그렇지만 두 번째 유형에서는 여러 필지에 한개의 건물이 존재하여 소유주가 임의대로 필지번호를 사용할 수 있기 때문에 어떠한 필지번호가 사용되었는지 알기 어렵다. 한편, 세 번째 유형은 한필지에 여러 개의 건물이 있어 동일한 지번에 여러 개의 대장자료가 있을 수 있다.



(그림 6-9) 필지번호와 건물의 관계 유형

나. 건물관련 행정자료 연계작업 과정

강영욱·이주일(2005)의 사례연구에서 건물도형과 건축물대장의 연계를 시도 했다. 그들이 다룬 사례지역은 서울시 창신동, 역삼동, 창천동이다. 각 동별 건축물대장(동별 개요)에 등록된 건축물과 건축물도형대장에 등록된 건축물은 각각 창신동이 3,039와 3,647개이며, 역삼동이 4,816과 5,415개이고, 창천동이 1,739와 1,796개이다.

건축물도형과 건축물대장간 연계 작업을 하기 전에 각 자료에서 연계키의 중복여부를 확인했다. 이는 건축물대장과 건축물도형의 연계키가 건축물기반이 아닌 필지기반이기 때문이며, 동일한 필지번호를 사용하는 건축물대장과 건축물도형이 존재하기 때문이다. 동일한 필지번호를 사용하는 건축물도형은 종로구 창신동이 34.6%, 강남구 역삼동이 19%, 서대문구 창천동이 20.5%이며, 건축물대장(동별 개요 자료를 기준)을 기준으로 볼 때는 종로구 창신동이 16.1%, 강남구 역삼동이 3.7%, 서대문구 창천동이 15.2%로 나타났다 (<표 6-8>).

또한, 자료 간 연계키가 서로 일치하지 않고 각 자료에만 존재하는 경우가 있어 이를 확인하였다. 건축물대장에는 있으나 건축물도형에는 해당 필지

번호가 없는 경우 종로구 창신동이 17.9%, 강남구 역삼동이 7.6%, 서대문구 창천동이 13.5%로 조사되었으며, 건축물도형은 있으나 건축물대장에는 해당 필지번호가 없는 경우는 종로구 창신동이 22%, 강남구 역삼동이 11.7%, 서대문구 창천동이 13.8%로 나타났다(<표 6-8>).

건축물대장 기준 연계율과 건축물도형 기준 연계율이 상이하며, 종로구 창신동과 강남구 역삼동의 경우 대장기준의 연계율과 도형기준의 연계율이 4%정도 차이가 난다. 건축물대장과 건축물도형의 평균 연계율은 종로구 창신동이 80%, 강남구 역삼동이 90%, 서대문구 창천동이 86%이다(<표 6-8>).

<표 6-8> 건축물 도형과 건축물대장의 연계

구분	종로구 창신동		강남구 역삼동		서대문구 창천동	
	해당 레코드수	비율	해당 레코드수	비율	해당 레코드수	비율
① 필지번호가 중복되는 건축물도형	1,263	34.6%	1,027	19%	369	20.5%
② 필지번호가 중복되는 건축물대장 (동별개요자료)	489	16.1%	177	3.7%	264	15.2%
③ 건축물대장에는 있으나 건축물도형에는 해당 필지번호가 없는 경우	544	17.9%	364	7.6%	235	13.5%
④ 건축물도형에는 있으나 건축물대장에는 해당 필지번호가 없는 경우	803	22%	634	11.7%	247	13.8%

건축물도형과 건축물대장 연계율	종로구 창신동	강남구 역삼동	서대문구 창신동
건축물대장 기준 연계율	82.1%	92.4%	86.5%
건물도형 기준 연계율	78%	88.3%	86.2%

필지번호 : 건축물도형을 식별할 수 있는 위치정보로 시/구/동 주소와 필지주소를 이용하여 작성한 19자리임

①, ④는 건축물도형 기준으로 비율 산정, ②, ③은 건축물대장 기준으로 비율 산정

건축물 도형과 건축물대장을 연계한 결과 세 개동(창신동, 역삼동, 창천동)의 도형기준 연계율이 각각 78%, 88.3%, 86.2%로 나타났다. 건축물도형을 기준으로 건축물대장 연계시 건축물대장의 동별 개요 자료의 필지번호를 이용하였으나, 건축물대장에는 해당 건축물의 대표지번이외에 관련지번 자료를 동별부속지번현황이라는 테이블에서 관리하고 있다.

건물도형기준 대장 연계율의 향상을 위하여 동별부속지번현황의 관련지번

을 활용하여 추가 연계작업을 수행함. 건축물도형과 건축물대장의 연계 방법은 다음 ①, ②와 (그림 6-10)과 같다.

① 1차 연계 : 건축물도형에 부여되어 있는 필지번호와 대장상의 필지번호의 자리수를 정비하여 연계함. 아파트와 같은 집합건축물 중 건축물명칭의 파악이 가능한 경우는 건물별로 연계함

- 건축물도형의 필지번호 : 편집지적도를 기반으로 생성한 번호로, 편집지적도가 갖는 시군구(5) + 법정동(5) + 대장구분(1) + 본번(4) + 부번(4)의 19자리로 구성되어 있음

<토지대장의 대장구분>

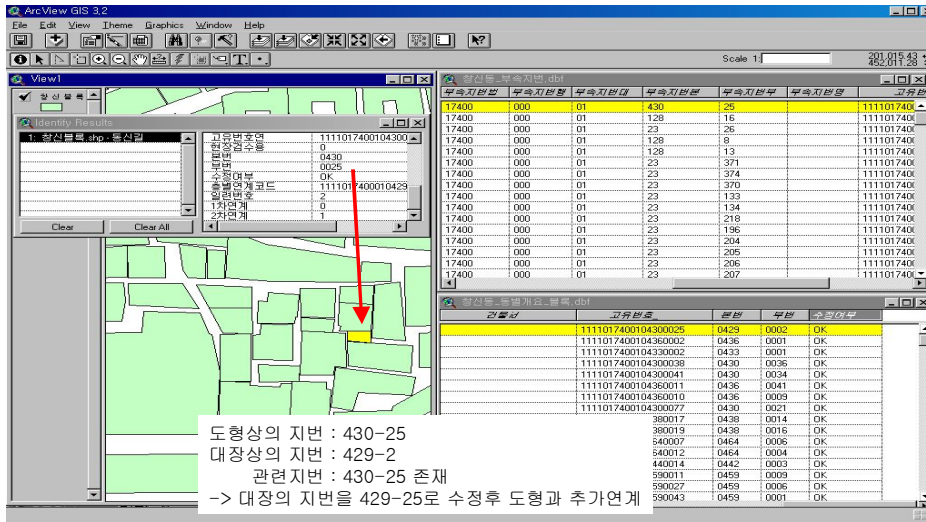
토지	임야
1	2

- 건축물대장의 필지번호 : 건축물대장 생성시 작성된 주소를 기반으로 작성한 번호로, 시군구(5) + 법정동(5) + 대장구분(2) + 본번(4) + 부번(4)의 20자리로 구성되어 있음

<건축물대장의 대장구분>

대지(토지)		산(임야)	
일반	집합	일반	집합
01	02	11	12

② 2차 연계 : 건축물대장의 필지번호와 건축물도형의 필지번호가 일치하지 않은 경우, 건축물대장의 부속지번현황 자료를 이용하여 건축물대장의 지번을 관련지번으로 수정하여 건축물도형과의 추가연계 작업을 수행함



(그림 6-10) 2차 연계 작업

건축물 대장의 부속지번현황자료를 이용하여 추가 연계작업한 결과 평균 5%이상의 연계율이 향상되었다(<표 6-9>, (그림 6-11) ~ (그림 6-16)).

<표 6-9> 건축물도형과 건축물대장의 연계율

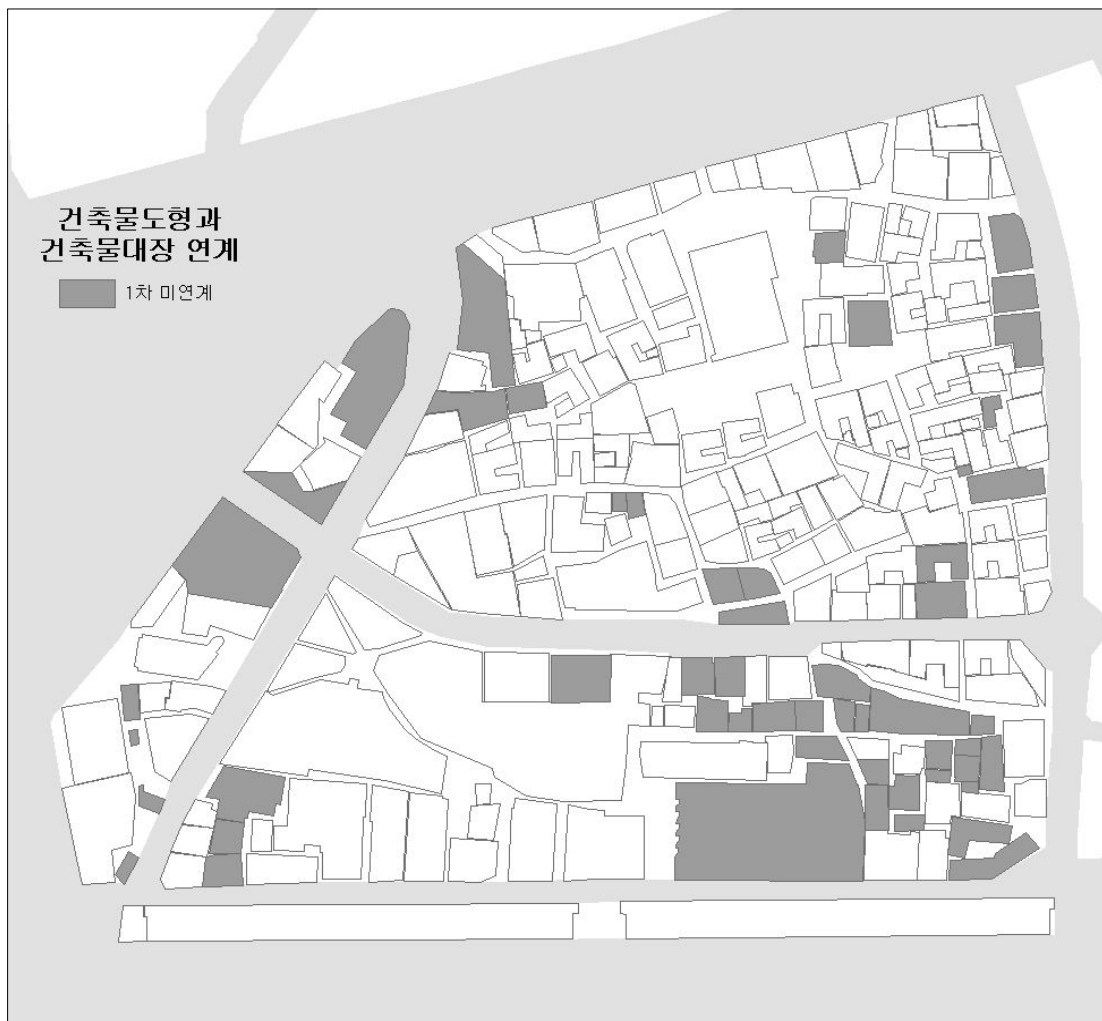
구분	총도형수	1차 연계		2차 연계		총연계	
창신동 블록건축물도형	242	188	77.7%	18	7.4%	206	85.1%
역삼동 블록건축물도형	266	220	82.7%	14	5.3%	234	88%
창천동 블록건축물도형	410	374	91.2%	21	5.1%	395	96.3%

다. 건물관련 행정자료 연계결과

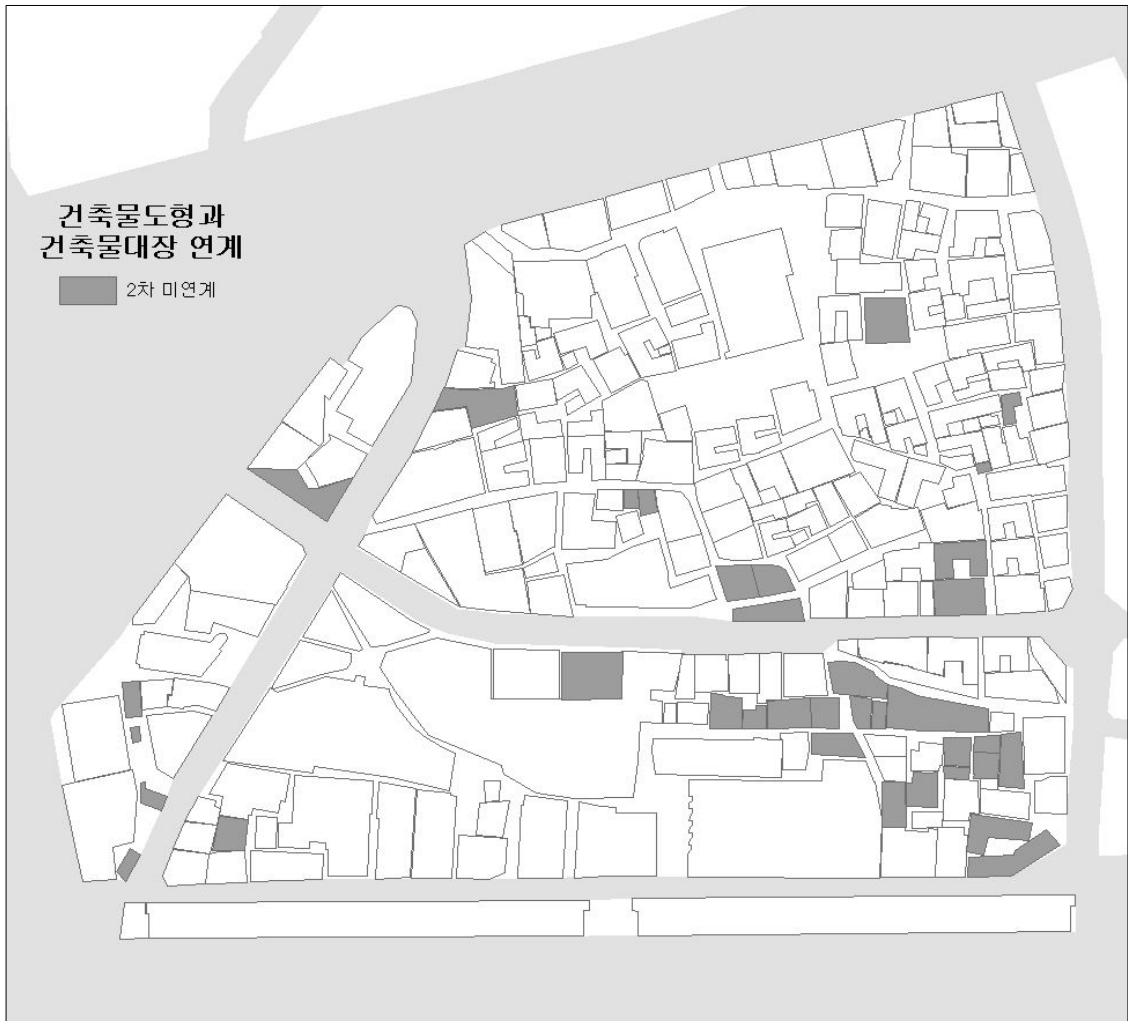
건물도형과 건축물 대장간 연계정도는 건축물대장을 기준으로 하여 82~92%, 건물도형을 기준으로 해서는 78~88% 정도까지 연계되는 것으로 나타났는데 이는 지번이 연계되는 것을 기준으로 한 것이며 연계결과에 대한 정확도를 나타내는 것은 아니다.

<표 6-10> 건물 도형과 건축물대장의 연계

건물도형과 건축물대장 연계율	종로구 창신동	강남구 역삼동	서대문구 창천동
건축물대장 기준 연계율	82.1%	92.4%	86.5%
건물도형 기준 연계율	78%	88.3%	86.2%



(그림 6-11) 건축물도형과 대장간 1차 연계 결과(창신동)



(그림 6-12) 건축물도형과 대장간 추가 연계 결과(창신동)



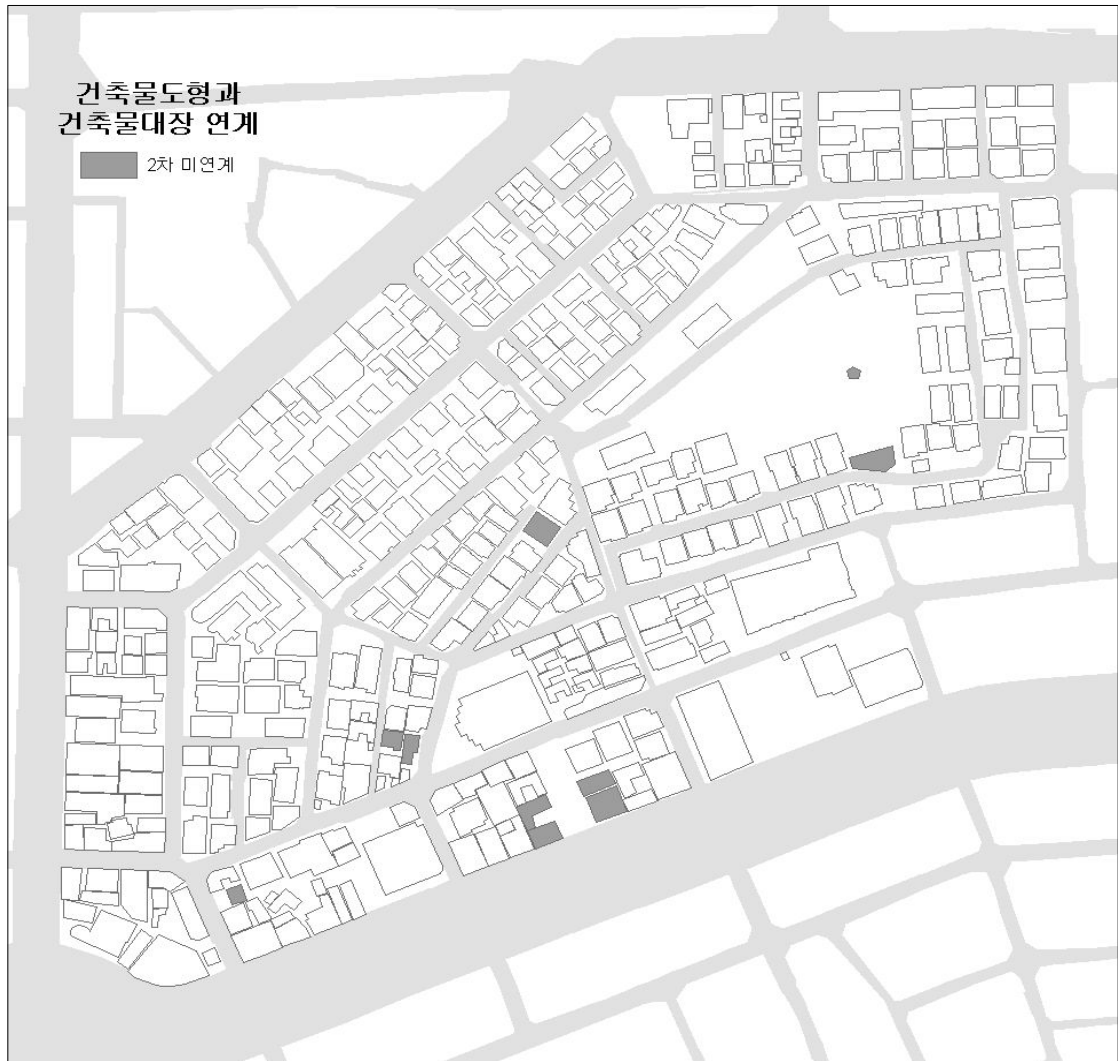
(그림 6-13) 건축물도형과 대장간의 1차 연계 결과 (역삼동)



(그림 6-14) 건축물도형과 대장간의 추가 연계 결과(역삼동)



(그림 6-15) 건축물도형과 대장간의 1차 연계 결과 (창천동)



(그림 6-16) 건축물도형과 대장간의 추가 연계 결과(창천동)

제4절. 행정자료 간 매칭 시 문제

1. 행정자료 간 연계의 문제

건축물 정보 정비사업에서는 건축물대장, 토지대장, 등기부등본을 연계한 바 있으며, 향후 건물도형, 과세대장, 위법건축물대장의 연계가 필요한 실정이다. 사례연구에서도 분석되었듯이 행정자료를 연계하는데 있어서 가장 큰 문제가 되는 것은 연계할 수 있는 연계키가 없다는 점이다. 즉 토지에 대한

정보는 필지번호를 가지고 고유키로서 연계가 가능하나, 건물에 대한 정보는 토지번호를 가지고 건물정보로서 사용하나 한 필지당 여러 개의 건물이 있기도 하고, 여러 개의 필지에 여러 개의 건물이 있다던가, 여러 필지에 한 개의 건물이 있어 대표지번을 무엇으로 쓰느냐가 행정자료들 간에 일치되어 있지 않아 건물 관련 행정자료들을 연계하여 활용하는데 많은 어려움이 있다.

또한 이외에도 자료 연계 시 발생하는 문제점은 다음과 같다.

- 행정구역의 불일치 : 행정구역이 변경되거나 명칭 변경시 건축물대장의 건축물 위치 정보에 이 사항이 반영되지 않은 경우
- 지번 불일치 : 토지이동에 따른 지번의 합병 및 분할 또는 고유번호의 변경이 있을 경우 건축물대장의 건축물 위치정보에 이 사항이 반영되지 않은 경우
- 물건명의 불일치 : 건축물의 물건정보는 건축물대장, 토지대장, 재산세, 종합토지세, 등기부등본, 국세청 기준시가 정보 등 다양한 부처에서 건축물의 물건명을 각각 관리하고 있어 물건명을 일치하기에 어려움이 있음

2. 행정자료의 정확도 문제

각각의 행정자료는 업무에서의 활용도에 따라 서로 다른 특성을 가지고 있으며, 각 자료의 정확도나 자료로서의 완전성이라는 측면에 있어 문제가 있다. 예를 들면 건축물대장의 경우 주요항목에 대한 누락률, 총괄표제부와 동별 현황 자료의 차이 등은 행정자료가 서로 연계되었다 하더라도 자료로서 활용가치를 상쇄하는 현상을 일으킨다. 또한 자료로서의 완전성뿐 아니라 과연 그 자료가 정확한 것인가 하는 것은 행정자료를 연계하였을 때 분석결과에 신뢰도에 영향을 미치는데 현재 행정자료들에 있어 정확도는 행정자료마다, 혹은 항목마다 다르기는 하지만 직접 연계하여 활용하기에는 많은 문제를 내포하고 있다.

3. 동일항목 비교 시 문제

행정자료 간 연계가 되었다 하더라도 동일항목에 대해 표준화된 분류체계를 가지고 있지 않으며, 활용상에 많은 제약이 따른다. 예를 들면 현재 분석 대상이 되고 있는 자료들을 각각 특성을 가지고 있어 어느 한 자료가 대표적으로 사용되기 보다는 상호보완적으로 사용이 필요하다. 예를 들어 어느 지역의 건축물이 활용되는 용도를 분석하고자 하는 경우 건축물 대장, 과세 대장, 개별특성조사자료 등을 연계한 후 공통되는 부분을 찾아내고, 나머지 부분은 각각의 행정자료 고유분을 분석하여 활용하여야 하는데 분류체계가 달라 분석이 어려운 문제가 있다.

다음은 건축물대장, 과세대장, 토지대장의 용도분류체계 및 건물구조에 대한 분류체계이다(<표 6-11>, <표 6-12>).

<표 6-11> 건축물 관련 자료의 구조분류체계 비교

건축물대장			과세대장		개별주택특성조사자료	
조적조	10	조적구조	11	철골콘크리트구조	01	철골(철골철근)콘크리트조
	11	벽돌구조	12	통나무조	02	통나무조
	12	블록구조	21	철근콘크리트조	03	철근콘크리트조
	13	석구조	22	철골조	04	철골조
	19	기타조적구조	23	석조	05	석조
콘크리트조	20	콘크리트구조	24	P.C조	06	P·C조
	21	철근콘크리트구조	25	목구조	07	목구조
	22	프리캐스트콘크리트구조	26	라멘조	08	라멘조
	29	기타콘크리트구조	27	스틸하우스	09	스틸하우스
철골조	30	철골구조	31	연와조	10	연와조
	31	일반철골구조	33	보강콘크리트조	11	보강콘크리트(보강블록조포함)
	32	경량철골구조	34	황토조	12	황토조
	33	강파이프구조	41	시멘트벽돌조	13	시멘트벽돌조
	39	기타강구조	51	목조	14	목조
철골철근콘크리트조	40	철골철근콘크리트구조	61	시멘트블럭조	15	시멘트블럭조
	41	철골콘크리트구조	62	경량철골조	16	경량철골조
	42	철골철근콘크리트구조	63	조립식판넬조	17	조립식판넬조
	49	기타철골철근콘크리트구조	71	석회및흙혼합벽돌조	18	석회 및 흙벽돌
목구조	50	목구조	72	흙벽돌조	19	돌담 및 토담조
	51	일반목구조	73	돌담및토담조	20	컨테이너건물
	52	통나무구조	74	컨테이너건물	99	기타
기타	99	기타구조	81	철파이프조		

<표 6-12> 건축물관련 자료의 용도분류체계 비교

건축물대장 (21단계)	재산세과세대장 (37단계)	개별주택특성조사 (주택부분만)	토지특성조사 (43단계)	토지대장 (24단계)
<ul style="list-style-type: none"> · 단독주택 · 공동주택 · 제1종 근린생활시설 · 제2종 근린생활시설 · 문화 및 집영 시설 · 판매 및 영업 시설 · 의료시설 · 교육연구 및 복지 시설 · 운동 시설 · 업무시설 · 숙박시설 · 위락시설 · 공장시설 · 창고시설 · 위험물 저장 및 처리 시설 · 자동차 관련시설(건설기계 관련시설 포함) · 동물 및 식물 관련 시설 · 분뇨 및 쓰레기 처리 시설 · 공공용시설 · 묘지 관련 시설 · 관광휴게시설 	<ul style="list-style-type: none"> · 전문하숙집 · 단독주택 · 다가구주택(공동주택) · 아파트 · 연립주택 · 다세대주택 · 사원아파트 · 여인숙 · 식품위생시설 · 콘도미니엄 · 호텔 · 대규모 점포시설 · 위락시설 · 의료시설 · 사무실 · 의식시설 · 위험물 저장시설 · 근린생활시설 · 문화시설 · 식품위생시설 · 유원시설 · 공중위생시설 · 숙박시설 · 교육연구시설 · 종교시설 · 체육시설 · 의료시설 · 시장(재래시장) · 생산시설 · 차량관련시설 · 운수시설 · 기타 · 농어가주택 · 광산주택 · 묘지 관련 시설 · 복지시설 · 농업생산시설 	<ul style="list-style-type: none"> · 단독주택 - 단독 - 일반 - 농가 - 광산 - 복지 - 별장 - 기타 · 다가구주택 - 다가구 - 일반 - 농가 - 광산 - 복지 - 별장 - 기타 · 다중주택 - 다중 - 기타 	<ul style="list-style-type: none"> 10. 주거용 -11. 단독 -12. 연립 -13. 다세대 -14. 아파트 -15. 주거나지 -16. 주거기타 20. 상업·업무용 -21. 상업용 -22. 업무용 -23. 상업나지 -24. 상업기타 30. 주상복합용 -31. 주상용 -32. 주상나지 -33. 주상기타 40. 공업용 -41. 공업용 -42. 공업나지 -43. 공업기타 50. 전 -51. 전 -52. 과수원 -53. 전기기타 60. 답 -61. 답 -62. 답기타 70. 임야 -71. 조림 -72. 자연림 -73. 토지임야 -74. 목장용지 -75. 임야기타 80. 특수필지 -81. 광천지 -82. 광업용지 -83. 염전 -84. 유원지 -85. 공원묘지 -86. 골프장 -87. 스키장 -88. 경마장 -89. 특수기타 90. 공공용지등 -91. 도로등 -92. 하천등 -93. 공원등 -94. 운동장등 -95. 주차장등 -96. 위험시설 -97. 유해, 혐오시설 -99. 기타 	<ul style="list-style-type: none"> · 전 · 답 · 과수원 · 목장용지 · 임야 · 광천지 · 염전 · 대 · 공장용지 · 학교용지 · 도로 · 철도용지 · 하천 · 제방 · 구거 · 유지 · 수도용지 · 공원 · 체육용지 · 유원지 · 종교용지 · 사적지 · 묘지 · 잡종지

4. 건축물 정보에 대한 총체적 현황 파악의 어려움

건축물대장은 건축과에서 적법한 인허가 절차를 거쳐 사용승인을 받은 건축물에 대해서만 대장이 작성된다. 인허가나 사용승인을 받지 못하여 발생한 무허가 건축물이나 불법 증개축과 용도변경과 같이 건축물을 사용하는 도중에 변경하는 사항에 대하여는 원칙적으로는 변경분을 신고하여 건축물대장에 관리하도록 되어 있으나, 실제로는 과세대장 및 무허가건축물대장에서 분리되어 개별적으로 관리되고 있는 실정이다. 적법, 위법, 가설 건축물 등 관리가 따로 되고 있어 실제 현황에 존재하는 건축물을 파악할 수 있는 총괄적인 대장이 없다.

- 건축물대장 : 사용승인을 취득한 적법한 건축물에 대하여 작성되는공부
- 기존무허가건축물 : 1981년 12월 31일 이전에 기존무허가건축물대장에 등재되었으나 철거보상이 완료되지 않아 무허가 건축물대장에서 별도로 관리하는 무허가 건축물
- 가설건축물
- 건축허가는 받았으나 건축물에 대한 사용승인이 나지 않은 경우 : 2000년도에 한시적으로 정리되었던 특정 건축물⁴³⁾ 제외

실세계에서 존재하는 건축물에 대한 총체적인 현황을 파악하기 위해서는 건축물대장, 과세대장, 위법건축물대장등을 모두 연계하여 살펴보아야 하는 어려움이 있다. <표 6-13>은 2004년 서울시 공간데이터웨어하우스에 저장되어 있는 자료를 기준으로 각기 대장별 건축물 수를 정리한 자료이다. 수치지형도의 건물수는 약 824,465동, 건축물대장은 722,890동, 과세대장은 655,761동으로 되어 수치지형도의 건물을 기준으로 볼 때 건축물 대장은 약 87.7%, 과세대장은 약 79.5%만이 대장으로 존재한다.

43) 특정건축물정리에 관한 특별법(2001)에 의해 건축허가를 받거나 신고를 하고 건축한 건축물 중 사용승인을 얻지 못한 건축물에 대하여 사용승인서를 교부함

<표 6-13> 서울시 대장별 건축물 수 비교

구별구분	건물도형		건축물대장		과세대장	
	건축물수	비율	건축물수	비율	건축물수	비율
종로구	34,907	4.2	30,912	4.3	26,679	4.1
중구	24,363	3.0	26,282	3.6	20,567	3.1
용산구	32,076	3.9	28,825	4.0	24,965	3.8
성동구	30,573	3.7	26,144	3.6	22,719	3.5
광진구	29,427	3.6	27,635	3.8	25,965	4.0
동대문구	45,380	5.5	37,672	5.2	33,769	5.1
중랑구	34,447	4.2	31,782	4.4	30,218	4.6
성북구	55,482	6.7	48,024	6.6	42,734	6.5
강북구	37,185	4.5	32,585	4.5	31,184	4.8
도봉구	21,548	2.6	16,922	2.3	16,328	2.5
노원구	21,062	2.6	14,828	2.1	13,392	2.0
은평구	50,069	6.1	45,563	6.3	40,293	6.1
서대문구	37,596	4.6	32,553	4.5	30,468	4.6
마포구	38,402	4.7	34,742	4.8	31,880	4.9
양천구	22,723	2.8	21,562	3.0	20,135	3.1
강서구	30,890	3.7	28,515	3.9	26,198	4.0
구로구	31,893	3.9	25,723	3.6	23,304	3.6
금천구	20,206	2.5	17,583	2.4	15,497	2.4
영등포구	40,630	4.9	34,538	4.8	29,555	4.5
동작구	36,168	4.4	32,368	4.5	28,662	4.4
관악구	41,579	5.0	35,663	4.9	34,033	5.2
서초구	25,430	3.1	19,873	2.7	18,094	2.8
강남구	28,263	3.4	23,792	3.3	22,700	3.5
송파구	29,266	3.5	25,650	3.5	24,471	3.7
강동구	24,900	3.0	23,154	3.2	21,951	3.3
총합계	824,465	100	722,890	100	655,761	100

2004년 서울시 공간데이터웨어 하우스 구축자료 기준

제5절. 시사점

본 절에서는 앞서 지적된 문제점 유형별로 어떠한 대안이 마련되어야 하는지를 정리해 보고자 한다.

1. 건축물 정보 간 연계의 문제

건축물 정보를 가지고 있는 주요 행정자료들 사용 시 가장 큰 문제는 연계할 수 있는 고유값이 없다는 점이었다. 이 부분은 2가지로 대안을 고려할

수 있는데, 단기적으로는 현재 건물, 토지, 등기 연계자료를 기반으로 건축물 대장과 재산세 과세대장, 무허가 건물관리대장, 수치지형도 건물에 대한 연계키를 부여하고, 행정자료의 소관부처로 하여금 이 지침을 따르도록 하는 것이다.

중·장기적으로는 현재 새주소부여사업의 결과 모든 건물에는 고유하게 건물번호가 부여되어 있으며, 국회의원 의원입법으로 법제화가 진행되고 있는 상황이다. 법제내용에는 주민등록을 비롯한 모든 공부상의 주소는 도로명 주소를 사용하도록 되어 있기 때문에 이에 따라 건물과 관련된 공부 및 수치지형도의 건물에는 도로명에 따른 건물번호를 부여하도록 한다.

2. 개별 자료들의 정확도 문제

연계키가 작성되었다 하더라도 개별 자료들의 정확도가 낮으면 분석된 결과에 대한 신뢰도가 떨어질 수 밖에 없다. 개별 자료의 정확도를 높이는 것은 행정자료를 관리하는 소관부처의 업무이지만, 각 소관부처는 고유의 목적에 따라 행정자료를 생성한 것이기 때문에 해당업무를 수행하는 데에는 문제가 없다고 할 수도 있다. 즉 이들 자료를 통계정보로서, 정책정보로서 활용하고자 할 때 문제가 발생하는 것이지 업무진행상으로는 큰 문제라고 인식하지 않을 수 있다는 점이다. 이들 문제를 해결하기 위해서는 개별 자료에 대한 특성을 명확하게 이해할 수 있고, 이들 데이터의 사용자가 특성에 따라 자료를 활용할 수 있도록 메타데이터를 만드는 것이 필수적인 일이 될 것이다.

즉 자료가 언제 생성되었으며, 유지관리는 어떻게 되고 있는지의 정보, 자료의 항목과 분류체계 정보, 이와 함께 항목별 누락률 등이 표기되어 있는 정보는 자료의 공유 및 공동 활용을 위해서는 필수적인 항목이다. 즉 현재로서는 행정자료들마다 전산화는 되어있지만 정확도가 낮은 문제를 한 번에 해결하기는 어려울 것으로 판단되며, 초기에는 메타데이터를 만들어 공유할 수 있는 환경을 조성하고, 차츰 각 행정자료들마다 정확도를 높여가는 작업을 하도록 해야 할 것이다.

3. 동일항목에 대해서도 비교 불가능한 문제

자료 간에 연계된 후에도 동일항목에 대해 분류체계가 상이한 것은 자료의 공동활용이나 정보로서의 분석에 걸림돌이 된다. 특히 용도, 구조 등은 건축물 대장, 과세대장, 등기부자료 등에서 사용하는 분류체계가 상이한데, 이러한 부분에 대해 표준분류체계를 작성하는 작업이 필요하다. 이는 다만 이 항목에서 뿐 아니라 가능한 모든 행정자료를 분야별로 유형화하고, 유형별로 자료항목을 비교분석한 후 동일항목 또는 유사항목에 대해서는 표준화된 분류체계의 작성이 이루어져야 하며, 표준분류체계 작성이후에는 각 소관 부처별로 이 분류체계를 따르도록 하는 일이 반드시 필요하다.

4. 건축물에 대한 총체적 현황 파악의 어려움

건축물관련 행정자료의 특성분석에도 보았듯이 각 행정자료들은 업무 특성별로 작성되어 실존하고 있는 건축물에 대한 총체적인 현황을 파악하기 어려운 실정이다. 이 문제에 대한 해결방법은 연계키를 작성하여 공통부분을 확인하고, 이 외에 각 행정자료별로 보유하고 있는 고유분을 확인하여 전체가 되도록 하는 것이 최선이라 판단된다. 즉 연계키를 작성하여 건물분야의 통합DB를 만들고 이를 기준으로 총체적 실체를 파악하도록 한다.

제7장. 순환총조사

제1절. 개요

인구주택총조사에서는 한 시점에 한 국가의 인구수나 주택수 등을 전수 조사한다. 이렇듯 일시에 전체를 조사함으로써 특정 시점에서 국가 전체의 상황과 더불어 국가 안에 있는 작은 지역의 상황까지 동시에 조망할 수 있는 장점이 있다. 특히 인구수와 주택수 등 국가 운영과 국민의 삶을 파악하는 가장 기본적인 요소로 이를 파악하여 국가의 기본 인프라에 대한 구체적인 상황을 한 눈에 알아볼 수 있다. 이러한 장점 때문에 대부분의 국가는 인구주택총조사를 주기적으로 실시하고 있다. 통상적인 총조사 주기는 5년 또는 10년이다.

전통적인 총조사가 갖는 명백한 장점에도 불구하고 현행 총조사는 몇 가지 현실적인 문제에 직면하고 있다. 첫 번째 문제는 총조사를 수행하는데 필요한 비용이 급속도로 증가하는 것이다. 우리나라 총조사 예산은 2000년에 834억 원으로 1인당 약 1,850원의 예산이 소요되었으나 2005년에는 1인당 2,860원으로 증가하여 약 55%가 증가하였다. 미국은 1960년에 1인당 소요예산이 \$2.5이었는데 1990년에는 \$10로 4배 증가하였다(Kish, 1990). 최근 노동비용은 급격히 증가하는 반면, 기술발전에 의한 비용절감은 이에 미치지 못하기 때문에 총조사 비용은 평균적인 경제성장률이나 물가상승률을 뛰어 넘어 급격히 증가하는 추세를 보이고 있다.

직면한 두 번째 문제는 총조사에서 수집하는 정보의 양이 증가하는 총조사 비용과는 반대로 도리어 감소하는 추세에 있다는 점이다. 1960년 미국 총조사에서는 전수조사에서는 7문항, 표본조사에서는 35문항을 조사하였으나 1990년 총조사에서는 전수조사 7문항, 표본조사 33문항으로 표본조사 문항수가 줄었다(Kish, 1990). 그리고 표본조사 비율이 1960년에는 25%, 1990년에는 16%임을 감안하면 총조사에서 수집되는 정보의 양은 줄어들었음이 명백하다.

세 번째로 제기되는 문제는 조사주기와 관련된 것이다. 총조사 주기가 5년이면 5년에 한번 씩 조사 결과가 갱신되기 때문에 시의성이 매우 떨어진다. 특히 요즈음과 같이 사회가 빠른 속도로 변하는 상황에서 5년 전 통계는 현실성이 떨어질 수 밖에 없다. 또한, 이전에 비해 지역사회의 역할이 더욱 중요해지면서 소지역 단위의 통계요구가 더욱 빈번해졌는데, 이러한 통계는 대규모 조사를 통해서만 확보할 수 있다. 이런 상황 변화에 따라, 시의성 높은 통계를 생산하기 위해서는 조사주기를 줄일 수 밖에 없게 되었다.

이러한 문제들을 해결하는 방안의 하나로, 여러 시점에 걸쳐 전체를 나누어 조사하는 순환총조사(rolling census)가 등장하였다. 순환총조사는 총조사가 안고 있는 문제가 부각되기 시작하면서 행정자료 활용과 함께 총조사에 대한 대안으로 심도있게 논의되기 시작하였다. 순환총조사에서는 전국을 겹치지 않도록 F등분하여 한해에 1/F씩 조사한다. 그리고 F기간 동안 조사한 것을 누적하면 전국을 총조사한 결과가 된다(Kish, 1990, 1998, 1999). 순환총조사의 장점은 무엇보다도 해마다 최신 자료가 생성되어 자료의 시의성이 개선되는 점이다. 또한 총조사 시점에 집중되는 예산 확보의 어려움을 여러 해로 분산하여 예산 확보의 어려움을 완화할 수 있는 장점이 있고, 해마다 조사를 실시하기 때문에 새로운 기술을 시험하고 개발하여 조사 방법을 개선하여 적용하기 쉬운 장점이 있다(ONS, 2003).

그러나 순환총조사는 현행 총조사가 지닌 한순간 전체를 조망하는 특성을 상실하는 단점이 있고, 현장조사에서 외부 영향이 증가할 위험이 있으며, 총조사 가치가 감소될 수 있는 약점이 있다. 또한 전체 조사비용을 분배함으로써 한해 조사비용은 작지만 이를 누적하면 전체 조사비용은 현행 총조사보다 증가하는 약점이 있다(ONS, 2003).

이론적인 관점에서 보면 순환총조사는 주기조사(periodic survey)의 일종이고, 주기조사의 일환으로 순환총조사 혹은 순환표본조사(rolling sample survey)가 연구되었다. 계속조사인 독립반복조사, 패널조사, 부분교체표본조사 등에 비하여 순환표본조사는 상대적으로 연구결과가 많이 알려져 있지 않다. Kish의 일련의 연구 결과와 이와 관련한 연구결과가 있을 뿐이다. 또한 순환조사가 적용된 사례도 그리 많지는 않다. 총조사로서는 프랑스의 순

환총조사가 유일하다. 순환표본조사의 대표적인 사례는 현재 실시중인 미국 지역사회조사(American Community Survey)이다. 미국지역사회조사는 향후 미국 총조사의 표본조사를 대체한다는 계획 아래 조사를 진행 중이다.

제2절. 외국의 사례

1. 프랑스 순환총조사

가. 도입 배경

프랑스에서는 1801년부터 2차 세계대전까지 5년마다 총조사가 정기적으로 시행되었다. 총조사의 목표는 모든 행정구역의 공식 인구를 결정하는 것과, 지역(commune) 수준부터 프랑스 전국에 이르기까지 모든 지역사회 단위의 인구와 사회적 특성을 제공하는 것이었다. 1999년까지는 이러한 틀에서 총조사가 시행되었다.

그러나 2차 세계대전 이후, 프랑스의 총조사는 일정한 주기가 없이 7년에서 9년 사이의 주기로 시행되어 왔다. 지난 30년을 살펴보면, 총조사 시행 연도는 1975년, 1982년, 1990년, 1999년으로 주기가 일정하지 않으며, 전전에 비해 길어졌다. 이렇게 주기가 일정하지 않은 데는 여러 가지 이유가 있지만 제도적인 문제에 기인하는 바가 크다. 프랑스 헌법에는 총조사 시행에 대한 명시가 없다. 단지 그때 그때 시행에 필요한 법령을 만들어 운영하고 있다. 총조사 시행에 가장 큰 걸림돌은 선거다. 프랑스는 3년, 5년, 7년을 주기로 각종 선거가 있는데 총조사는 가능하면 이를 피해 시행된다. 이 때문에 일정한 주기를 갖지 못하게 되었다. 이에 더하여 최근에는 예산의 압박 또한 그나마 불규칙적인 조사 주기를 더욱 늘어지게 하는 데 큰 몫을 했다. 특히 1999년 총조사는 예산상의 제약으로 원래는 1997년에 시행했어야 하는 것을 연기한 것이다(Durr and Dumais, 2002; INSEE 2004, 2006).

국가적으로는 조사주기가 늘어나는 상황에서 지방정부의 요구는 전혀 반대의 방향으로 진행되고 있다. 프랑스에서는 20년 이상 분권화가 진행되어

지방정부에 더 많은 권한이 주어졌고, 이에 따라 지방정부는 정책수행에 필요한 새로운 지역정보가 더 많이 필요하게 되었다. 따라서 지역정보의 주된 공급원인 총조사에 대한 보완이 요구되었다. 이와 더불어, 지방정부가 받는 교부금은 총조사 인구에 근거를 두고 있기 때문에, 최근 성장하는 지방정부는 총조사 중간에 중간조사를 통해서라도 인구수를 자주 집계해야 한다고 주장하였다(ONS, 2003).

다른 이유는 중앙통계국(INSEE)과 자치구(commune)의 역할에 관한 것이다. 자치구는 프랑스에서 최소단위 행정구역으로, 자치구의 장은 독립된 자치구를 대표하며 동시에 주(state)를 대표한다. 총조사에서 자치구장은 총조사에 필요한 조사원 고용에 개입하지만, 고용된 조사원은 중앙통계기관에 의해 관리되었다. 따라서 총조사에서 중앙통계국과 자치구의 역할을 분명히 할 필요가 있었다.

마지막으로, 8-9년마다 시행되는 총조사는 중앙통계국과 자치구에게 다음 총조사에서는 사용할 수 없는 자료처리 프로그램 등과 같은 일회성 투자가 커다란 부담이 되었다. 또한 대도시에서는 증가하는 무응답률로 조사가 점점 어려워지고 있다. 많은 수의 조사원들을 관리하고 높은 품질의 자료를 수집해야 하는 중앙통계국으로서는 이러한 어려움을 시간에 따라 분배하여 줄이는 방안을 모색하게 되었다.

이러한 이유 때문에 중앙통계국은 90년대 후반에 총조사에 대한 재설계 프로젝트를 구상하였다. 제약사항은 관련예산을 증액하지 않는다는 것이었다. 1995년에 총조사 개혁의 기본 방향이 제시되었고 2001년에는 2004년부터 실시할 새로운 총조사를 수행할 팀이 만들어졌다. 프랑스 순환총조사는 2004년에 처음으로 도입되어 실시되었다(Durr and Dumais, 2002; INSEE, 2004, 2006).

나. 조사방법

(1) 순환총조사 원칙

새로운 방법의 첫 번째 원칙은 중간년도를 위해 5년 순환주기 동안 매년 자료를 생산하는 것이다. Y-4, Y-3, Y-2, Y-1년에 수집된 자료와 Y년의 자료의 “평균”이 Y-2년의 대표값으로 사용된다. 두 번째 원칙은 예산의 제약 아래 표본조사를 수행하는 것이다. 총조사 주기가 7년이 적절하다고 간주할 때, 한해 예산은 총조사 예산의 1/7이 된다. 결과적으로 5년에 걸쳐 총조사 예산의 5/7을 사용하고 인구의 5/7를 조사하며, 한해에는 총조사 예산의 1/7로 총인구의 1/7을 조사하게 된다.

(2) 표본추출방법

프랑스에는 약 37,000개에 달하는 자치구가 있고, 각 자치구 크기가 매우 다르다(<표 7-1> 참조). 크기 만 명 미만의 소규모 자치구 수는 3만 5천개가 넘고, 포함하는 인구비율은 약 50%이다. 반면 만 명 이상의 중대규모 자치구는 그 수는 많지 않으나 인구 비율은 약 50%이다. 따라서 소규모 자치구는 5년 순환주기로 전수조사가 고려되었고, 중대규모 자치구는 표본조사가 고려되었다.

<표 7-1> 프랑스 자치구 현황

	소규모 자치구	중대규모 자치구	합계
	만 명 미만	만 명 이상	
자치구의 수	약 35,750개	약 1,250개	약 37,000개
인구 수	약 50%	약 50%	100%
표집 비율	100%	40%	
표집된인구수	50%	20%	70%

(3) 소규모 자치구

소규모 자치구는 5개 그룹으로 분할되었다. 5개 그룹은 약 10개 기준에 의하여 균형이 맞도록 하였는데 그 기준은 다음과 같다: 거주지 수(# of dwellings), 집단건물(collective buildings)의 거주지 수, 나이 그룹(20미만, 20-39, 40-59, 60-74, 75이상), 성별(남, 여). 전국 수준과 26개 프랑스 지역(region) 수준에서 균형이 맞도록 하였으며, 방법론으로는 균형 표본(Balanced samples, Deville-Tille)을 사용하였다. 매년 한 개 그룹에 속한 모든 자치구의 거주자가 모두 집계되고, 5년에 걸쳐 소규모 자치구의 모든 거주자가 조사된다.

(4) 중대규모 자치구

각 중대규모 자치구에서는 해마다 8%의 인구가 표본조사 된다. 표본추출을 위한 프레임으로 건물일람표(RIL : Inventory of Located Buildings)가 1999년 총조사 자료를 기초로 하여 최초로 작성되었다. 건물일람표는 모든 건물의 목록으로 건물은 주거용, 행정용, 산업용, 상업용으로 식별되어 GIS로 위치가 파악된다. 건물일람표는 건물허가, 지방세 기록, 우편 번호 등의 행정자료를 이용하여 유지 관리하는데 해마다 자치구에서 점검하고 중앙통계국에서 최종 검토한다.

자치구 안에서 대규모 주소지(addresses)와 신축 거주지는 전수조사된다. 나머지 지역은 인구통계적 특성에 따라 균등하게 분포하도록 5개 균형그룹으로 나뉘어 한해에 한 그룹에서 약 40%의 거주지가 표본조사된다. 이렇게 하면 해마다 각 자치구 안에서 약 8%의 거주지가 조사되고 5년에 걸쳐 중대규모내 거주지 중 약 40%의 거주지가 표본조사된다.

결과적으로 해마다 약 8,000개 자치구, 400만개의 거주지, 820만명의 거주자가 조사되며, 5년 주기 동안 약 37,000개 자치구, 4,250만 거주자(총 인구의 약 70%)가 조사된다.

(5) 자료수집 방법

자료 수집 방법은 호별 가구방문, 배달, 집계 등이 혼용된다. 질문지는 가구용과 개인용 2종류인데, 가구용에서는 거주자 리스트와 거주지 전형에 따른 특성, 자동차 수 등에 관한 15개 문항이 있고, 개인용 질문지에는 나이, 성별, 출생지, 국적과 이전 5년의 거주 지역, 그리고 학력, 직업, 직장 등 25개 문항이 있다.

다. 추정 방법

순환총조사의 첫 번째 주기가 끝나는 2008년부터는 3종류의 결과들이 매년 생산된다. 마지막 총조사 시행년도와 무관하게 모든 자치구의 공식인구가 발표된다. 이때 모든 자치구의 공식인구수는 (Y-2)년의 인구수로 발표한다. 또한 지난 5년 동안에 수집된 자료의 결합에 근거하여 중간년도의 모든 지리적 수준에서의 인구 자료, 그리고 당해 연도 조사에 근거한 프랑스 전국과 지역(region)에 대한 자료가 발표된다.

한 해의 대표값으로서 공식인구(official populations)를 추정하는 방법은 자치구의 크기에 따라 다르다. 만 명 이상의 중대규모 자치구에서는 최근 5년 동안 집계된 자료를 합계하여 중간년도의 건물일람표에 따라 가중합으로 평균인구가 추정된다. 반면 소규모 자치구에서는 중간년도의 인구수를 이용하여 이전 연도 인구수는 세금자료 등 외부자료를 이용하여 외삽(extrapolation)하고, 이후 연도 인구수는 전년도 자료를 이용하여 내삽(interpolation)한다.

라. 조사 조직

(1) 시기별 업무

연중 진행되는 순환총조사의 시기별 업무는 다음과 같다. Y년 총조사를

위하여 (Y-1)년 1월부터 6월까지는 건물일람표를 갱신하고, 7월부터 8월까지 표본을 선정한다. 7월과 12월 사이에 선정된 자치구를 출력하고 배포하며, 9월과 12월 사이에는 자치구와 감독관을 훈련하고 총조사를 홍보한다. 12월에는 조사원을 채용하고, Y년 1월에는 채용된 조사원을 훈련한다. 그리고 Y년 1월 3번째 목요일부터 조사를 시작하여 5주 동안 계속 조사를 진행한다. 조사된 데이터의 집계, 가공은 3월부터 시작하여 12월 까지 진행한다. 첫 번째 결과는 12월에 발표된다.

(2) 필요 인력

순환총조사를 수행하는데 한해에 필요한 인력은 중앙통계국에서 총조사 진행 중앙 담당자 30명, 컴퓨터 전문가 20명, 지역 사무소 담당자 200명, 건물일람표 갱신을 위한 지역 담당자 100명, 자치구를 전담하는 감독관 550명 등이다. 또한 자치구에서는 감독관 8,000명, 조사원 18,000명이 필요하다.

(3) 소요 예산

한해에 필요한 소요 예산은 총 2,900만€인데, 항목별로 보면, 자치구 비용 배상금 1,800만€, 데이터 획득 비용 450만€, 양식 출력 비용 150만€, 홍보 비용 100만€, 중앙통계국 감독관 비용 110만€이다.

(4) 중간 평가

2004년과 2005년 총조사 실시 결과를 평가하였는데, 자료 수집에는 큰 문제는 발생하지 않았다. 2004년 소규모와 중대규모 100개 자치구 인구수는 2005년 1월에 활용 가능하며, 해당 자치구의 인구통계적, 사회적 특성은 2005년 7월에 활용 가능하다. 그리고 2005년 소규모와 중대규모 200개 자치구 인구수는 2006년 1월에 활용 가능하고, 해당 자치구의 인구통계적, 사회적 특성은 2005년 7월에 활용 가능하다.

조사원이 접촉을 못했거나 응답거부 비율은 2004년에는 평균 3.3%, 였고, 2005년에는 평균은 2.8%였다. 무응답은 대체(imputation)하여 처리됐다.

<표 7-2> 무응답 비율

	2004년	2005년
만명 이상 자치구	5.8%	4.9%
만명 미만 자치구	2.2%	2.0%
합계	3.3%	2.8%

마. 성공 요인

순환총조사가 성공하기 위한 요인으로는 (i) 양질의 표집 프레임, (ii) 자치구와 파트너십이 지적되고 있고, 추가 요인으로 (iii) 행정자료 가용성(availability)이 언급되고 있다. 양질의 표집 프레임을 구축하기 위해서는 프레임 갱신에 행정자료 파일이 활용되어야 하고, 행정자료 활용에 대한 규정에 제정되어야 한다. 또한 자치구와의 파트너십을 원활히 하기 위해서는 지역 수준에서 행정자료 파일이 적절하게 점검될 수 있어야 한다.

2. 미국지역사회조사

가. 도입 배경

미국 총조사는 전수조사와 표본조사를 병행하여 10년 주기로 실시된다. 전수조사에서는 인구조사에서 필수적인 최소한의 기본 질문으로 두 개의 가구 항목(가구원수, 가구명)과 6개의 개인항목(이름, 성별, 연령, 응답자와의 관계, 히스패닉계 여부, 인종)만을 조사한다. 표본조사는 전국 가구의 1/6을 대상으로 전수조사 항목을 포함한 가구 항목 2개와 개인 항목 53개 문항을 조사한다. 표본조사는 각 주(state)를 비롯하여 대도시나 대규모 카운티(county) 및

소규모 카운티의 인구, 경제적 통계자료를 제공한다(Alexander, 2000, 2002).

총조사가 제공하는 지방 단위 및 부분모집단(subpopulation)수준의 상세한 인구정보는 많지 않다. 미국은 해마다 100만 가구, 200만 명 이상의 인구가 증가하는 등 급속히 변화하기 때문에, 해마다 새로운 자료에 대한 요구가 많다. 그러나 총조사는 10년 주기로 인하여 시의성이 떨어지며, 표본조사로는 소지역까지의 상세한 자료를 제공하지 못하는 한계점을 가지고 있다. 소지역 단위의 신뢰도 높은 자료의 필요성이 계속적으로 지적되었다.

미국 센서스국(Census Bureau)에서는 1992년 이후 표본조사를 대체하는 과정에서 연속측정(continuous measurement)를 2000년 총조사의 연구과제로 상정했고, 그후 연속측정은 미국지역사회조사(American Community Survey, ACS)로 개명하였다. 센서스국에서는 2010년 총조사에서 표본조사를 미국지역사회조사로 대체할 계획을 가지고 있다.

미국지역사회조사가 도입된 이유는 다음과 같다.

- 10년 동안 표본조사와 유사하도록 지리적으로 상세한 경제사회자료를 제공한다.
- 미국지역사회조사는 매월 행해지는 조사이므로 직원의 대다수를 종신고용으로 한다. 이를 위해서 임시직원에게 크게 의존하여 10년 간격으로 행해지는 총조사보다 높은 품질의 결과를 기대할 수 있다. 또한 데이터수집에 대한 지속적인 개량이 가능해진다.
- 총조사로부터 표본조사의 부분을 삭제함으로써, 운영을 간소화하고 인구 조사와 포함범위(coverage)의 향상에 집중할 수가 있다.
- 주(state)나 지방자치단체의 복지개혁 등에 대한 새로운 책무를 지원한다.
- 미국지역사회조사의 결과를 전수조사 및 다른 연방정부의 가구조사 및 행정자료 등과 통합하여, 연방정부통계시스템의 인프라를 개선한다.

나. 조사 방법

(1) 표본설계(sample design)

미국지역사회조사에서 채택한 순환표본 추출법은 일단계 계통추출법(unclusterd one-stage systematic sampling)으로 중복하지 않고 매월 새로운 표본을 뽑는다. 중복성이 없으므로 가법성이 유지되고 추계치를 시간 및 지리적 범위를 넘어선 다양한 조합으로 누적할 수가 있다. 표본의 크기는 25만 주소지(address)로 매월 새로운 표본으로 교체된다. 연간 표본의 수는 약 300만이다. 이 표본설계에서는 동일 가구가 다시 샘플로 뽑히기까지는 적어도 5년이 지나야 된다.

표집 비율(sampling rate)은 집계 단위(governmental unit)의 크기에 따라 다르다. 집계 단위의 크기가 200가구 이하이면 추출률은 10%(5년간 50%)에 이르지만, 2000가구 이상인 단위에서는 추출률이 1.7%(5년간 8.5%) 이하이다(<표 7-3> 참조). 표본은 소지역에서 총조사의 표본조사를 대체하기 위하여 5년간 누적된다.

<표 7-3> 표본 추출률

집계 단위 크기(governmental unit size) : 포함된 주택(occupied housing unit) 수	표본 추출률	
	2000년 총조사	ACS 추출률 (5년 누적)
0-200	50.0%	50.0%
201-800	50.0%	~ 35.0%
801-1200	25.0%	~ 17.5%
1201-2000	16.7%	~ 12.0%
2000 초과	12.5%	~ 8.5%

(2) 표본추출 프레임

미국지역사회조사는 표본추출 프레임으로 갱신된 마스터 주소 파일

(undated Master Address File, MAF)을 사용한다. 센서스국은 미국 우편공사의 주소록 및 지방자치체의 주소록을 사용하여 1990년 총조사에서 사용된 주소 관리 파일(Address Control File, ACF)를 갱신한 후 TIGER라고 하는 센서스국의 지리DB에 연계시켜 MAF를 개발하였다.

(3) 자료 수집

자료 수집(data collection)은 총조사와 인구조사의 경험을 기초로 우편조사, 전화조사, 대면조사(personal visit)를 병행한다. 우편조사는 표본 전체를 대상으로 하며 우편조사의 응답률을 높이기 위하여 복수로 질문지를 발송한다. 우편 응답이 불완전한 경우는 전화조사로 추적조사(follow-up)를 한다. 컴퓨터보조-전화조사(Computer-assisted telephone interview, CATI) 방법으로 통화를 위하여 3회를 시도하며 전화조사는 4주간 계속된다. 전화조사도 영어와 스페인어로 진행한다. 대면조사는 컴퓨터를 휴대하는 컴퓨터보조-대면조사(Computer-assisted person interview, CAPI)를 실시하는데 대면조사의 대상은 우편조사와 전화조사에서 응답하지 않는 표본가구와 우편조사 중 불성실 응답을 한 응답자이다.

2002년 조사에서는 우편조사로 58.9%, 대면조사로 31.6%, 그리고 전화조사로 9.5%가 응답을 받았다. 미국지역사회조사에서의 단위무응답(unit nonresponse) 비율은 모든 주에서 매우 낮게 나타나며, 전국 수준에서 2000년-2005년 조사에서 모두 95%이상의 응답률을 보였다.

미국지역사회조사는 영어로 진행하는 것이 기본이나 스페인어 조사도 병행한다. 우편조사는 영어 질문지가 기본이나 요청에 의하여 스페인어 질문지가 발송되고, 전화조사와 대면조사에는 영어와 스페인어를 동시에 사용하는 면접자가 면접을 진행한다. 언어 문제와 함께 중요한 문제는 응답률이다. 우편조사에서 높은 응답을 받는 것이 무엇보다 중요하며, 전화조사 및 대면조사에서 응답률을 높이는 효과적인 방법을 찾는 것이 매우 중요하다.

(4) 자료 처리

자료 처리(data process)는 1년간 데이터를 누적하는 것을 기본으로 한다. 응답자와 무응답자를 구별하고 무응답의 비일치성(inconsistent)를 식별한다. 그리고 항목 무응답은 대체(imputation) 처리를 하는데, 총조사에서 사용한 방법과 동일한 방법을 사용한다. 규칙에 근거한 대입(rule-based assignment)과 최근방 할당(nearest neighbor allocation)을 한다.

무응답률이 낮은 항목은 성별, 시민권(citizenship), 전화 가용성(telephone availability), 월세(monthly condominium fee), 동거하는 손자녀 수 등이고, 무응답률이 높은 항목은 연간 휴대전화 사용료, 연간 주택 보험료(yearly property insurance), 월간 주택담보 지급료(monthly mortgage payments), 연간 부동산세(yearly real estate taxes) 등이다.

(5) 가중과 추정

연간 가중치 보정은 3단계 절차를 거친다. 초기 가중치는 표본추출확률을 반영한다. 표본추출확률의 역수가 초기 가중치가 된다. 다음으로 응답가구와 무응답 가구 비율을 가중치에 반영한다. 응답률의 역수를 가중치에 곱하여 가중치를 보정한다. 마지막으로 성별, 나이, 인종 등 인구 통계적 모집단 특성치 비율을 반영하여 가중치를 보정한다.

대부분의 항목에서 다년 추정치(multi-year estimates)는 매년 추정치의 평균으로 계산된다. 현재 고려되는 다년 추정치는 3년 추정치와 5년 추정치이다. 그런데 다년 추정치와 매년 추정치는 차이가 있기 때문에 이에 대한 심층 연구를 필요로 한다. 예로서 특성치의 값이 해마다 증가 추세에 있는 항목과 해마다 증가와 감소를 반복하는 항목을 고려하자.

- 예제 1 : 해마다 증가 추세에 있는 항목으로 해외에서 태어난 인구 비율을 고려하자. 어느 지역의 해외 출생 인구 비율의 일년 추정치가 다음과 같다고 하자. 그리고 3년 추정치와 5년 추정치가 일년 추정치의 평균으로 구

한다고 하자(편의상 연도별 가중치는 고려하지 않았음). 그 결과는 아래 <표 7-4>와 같다. 서로 다른 추정치를 어떻게 활용할 것인가?

<표 7-4> 일년 추정치와 다년 추정치 : 증가 추세 항목

연도	2005	2006	2007	2008	2009	2010
1년 추정치	20.0%	21.2%	23.3%	28.6%	32.6%	35.1%
3개년 추정치(2005-2007)	21.6%					
3개년 추정치(2006-2008)		24.4%				
3개년 추정치(2007-2009)			28.2%			
5개년 추정치(2005-2009)	25.1%					
3개년 추정치(2008-2010)				32.1%		
5개년 추정치(2006-2010)	28.2%					

- 예제 2 : 해마다 등락을 거듭하는 항목으로 주택 소유 비율을 고려하자. 아래 <표 7-5>는 어느 지역의 주택 소유 비율의 일년 추정치와 다년 추정치이다. 3년 추정치와 5년 추정치는 일년 추정치의 평균으로 구하였다.

<표 7-5> 일년 추정치와 다년 추정치 : 등락이 있는 항목

연도	2005	2006	2007	2008	2009	2010
1년 추정치	52.0%	66.7%	44.3%	40.0%	55.6%	61.5%
3개년 추정치(2005-2007)	54.3%					
3개년 추정치(2006-2008)		50.3%				
3개년 추정치(2007-2009)			46.6%			
5개년 추정치(2005-2009)	51.7%					
3개년 추정치(2008-2010)				52.4%		
5개년 추정치(2006-2010)	53.6%					

위의 두 가지 예에서 보듯이 만일 여러 종류의 추정치, 예를 들면 일년 추정치, 3년 추정치, 5년 추정치, 가 동시에 발표되면 이용자는 어떤 추정치를 어떻게 이용해야 하는가? 통계 제공자는 이를 어떻게 해석해야 하는가? 추세 분석(trend analysis)과 더불어 제공하는 추정치에 대한 심층적인 검토가 요구된다.

(6) 결과 발표

2006년부터는 연간 추정 결과를 매년 발표하고, 2008년부터는 3개년 추정치를 매년 발표하며, 2010년부터는 5개년 추정치를 매년 발표할 계획이다.

다. 몇 가지 중간 결과

2004년 센서스 통계국은 1999년, 2000년, 2001년 3개년의 ASC 추정 결과와 2000년 센서스의 결과를 비교하였다. 몇 가지 중요한 발견을 살펴보면 다음과 같다(U.S. Census Bureau, 2004).

- 카운티 수준에서는 대부분의 ASC 추정치와 센서스 추정치가 같았다.
- 일부 카운티 수준의 차이는 1%p 이하로 작았다.
- 일부 카운티 수준의 차이는 설문지나 자료 정리, 조사방식의 차이 때문이었다.
- 센서스트랙 수준의 차이는 작은 표본크기와 큰 분산으로 찾기 어려웠다.
- 일부 변수에서 두 조사의 통계적으로 의미 있는 차이가 나타난 카운티가 많았다. 9개 또는 그 이상의 카운티에서 통계적으로 의미 있는 차이를 나타낸 변수는 16개였다. 그러한 변수는 장애상태, 인종, 방의 수다.

라. 남은 과제

2005년부터 연 300만 표본을 조사하기 시작하였고, 2006년부터는 완전 표본(full sample)에 의한 첫 번째 결과물이 나오고 있다. 가구조사의 계속되는 문제점은 미국지역사회조사에서도 그대로 남아 있으며 지속적으로 개선을 해야 한다.

미국지역사회조사에서 발생하는 특수한 문제로는 마스터 주소 파일 갱신 문제, 일년 추정치와 다년 추정치의 심층 연구 및 이용자 안내, 그리고 대표본(년간 300만 표본) 조사의 품질 유지 문제 등이 있다.

- 마스터 주소 파일의 완전도 : 미국지역사회조사의 표본 프레임은 MAF이

다. 해마다 새로 생긴 가구를 정확하고 신속하게 MAF에 도입해야 하고, 없어진 주소를 MAF에서 제거해야 한다. 또한 번지, 도로명에 의한 전형적인 도시주소가 아닌 주소를 어떻게 수집, 등록할 것인가도 과제이다.

- 일년 추정치 및 다년 추정치 활용 : 2010년부터는 일 년 추정치, 3개년 추정치, 5개년 추정치가 동시에 발표된다. 이에 대한 비교 연구 및 이용자 안내가 추가 되어야 한다.
- 표본 크기가 연간 300만인 조사는 매우 규모가 큰 조사이다. 데이터 품질을 유지하는 문제가 지적되지 않을 수 없다. 데이터의 정확성 못지않게 발표 결과의 시의성, 비교성, 일치성 등이 높은 수준을 유지해야 한다.

제3절. 순환총조사 도입 시 검토사항

순환총조사는 전통적인 총조사와는 매우 다른 형태의 총조사이다. 현행 총조사에 대한 예산, 수집하는 정보량 문제 등으로 순환총조사가 대안으로 검토되고 있으나 순환총조사는 나름대로의 장단점이 있기 때문에 순환총조사를 실제로 도입하기 위해서는 이에 대한 준비를 충분히 하여야 한다.

통상적인 순환총조사의 장점은 총조사 결과가 매년 생산되어 자료의 시의성이 개선되고, 해마다 조사하기 때문에 조사 항목을 좀 더 유연하게 할 수 있으며, 매년 조사하기 때문에 조사 방법을 계속하여 개선할 수 있다. 그리고 총조사에서 고용하는 인력과 지원 장비를 계속하여 사용할 수 있기 때문에 조사 인력 및 예산을 효율적으로 운용할 수 있는 장점이 있다.

단점은 총조사의 가장 큰 명분인 ‘한 순간에 전체를 조망하는 특징’이 사라지고, 이 때문에 국가적인 지원을 받던 ‘총조사’의 중요성이 희석될 수 있는 점이다. 또한 현장 조사에서 외부로부터의 영향이 개입될 위험성이 증가할 수 있고, 통계 이용자들의 총조사에 대한 신뢰가 줄어들 수 있으며, 총조사의 총 예산이 증가할 수도 있다.

이런 장점과 단점을 염두에 두고, 아래는 순환총조사를 도입할 때에 고려해야 하는 사항을 구체적으로 살펴본다(ONS, 2003).

1. 총조사 결과물

순환총조사의 큰 장점은 현행 총조사가 제공하는 정보들을 더 자주 제공하여 인구통계를 발전시킬 수 있다는 것이다. 그러나 이 방법은 매년 전국의 일부분에 대한 자료만을 수집함으로써 균형을 맞추기 때문에, 전 인구에 대한 통계를 한 시점에 제공하지 못하고 여러 다른 시점에 조사하고 결합하여 제공함으로써 지역 간의 비교가 복잡해지는 단점이 있다. 또한 총조사는 통근·통학 통계와 같이 지역 간 이동과 관련된 데이터도 수집하는데, 이러한 이동 자료에 대한 총조사의 영향을 평가하기 위해서는 추가적인 연구가 필요하다.

한순간에 전국을 조망하는 특성의 손실을 보충하기 위하여, 표본에 포함되지 않은 지역들에 관한 모수 추정을 고려해야 하고, 수집된 자료를 사용하는 여러 가지 방법을 살펴보아야 한다. 소지역에 대한 모수 추정을 하거나, 모수 추정값과 일치하는 개인 단위의 데이터베이스를 완전히 모델링 혹은 대체하는 것도 가능할 것이다. 중요한 문제는 이러한 추정값이 전통적인 방법에 의해 총조사 주기 중간 연도에 생산되는 현재의 것보다 얼마나 더 좋을 것 인가이고, 데이터를 어느 정도까지 모델링할 수 있게 데이터베이스를 만들 수 있는가 하는 점이다. 이러한 문제들은 조사 설계 전략(design strategies)와 깊은 연관이 있기 때문에 그와 관련하여 다뤄야 한다. 모든 결과 데이터의 특징은 그것들이 모두 소표본에 근거한 추정치라는 것인데, 이 특징은 전통적인 센서스 사용자들에게 불만이 될 수도 있다.

순환총조사는 현행 총조사 항목을 모두 포함할 수 있으며 조사 기간 동안 문제의 내용을 바꿀 수도 있다. 바꾸는 내용의 폭이 클 경우, 연속된 기간 사이에는 비교가능성이 줄어들겠지만, 총조사 결과를 변화하는 사용자의 요구에 맞춰 제공하는 결과가 된다.

2. 사용자 수용성

인구조사기관(OPCS, 1992)에 따르면 지방정부들이 순환총조사를 반대하

는 일반적인 이유는 지역 간 비교가능성을 잃기 때문이다. 그러나 순환총조사에는 가능한 대안적인 설계가 많이 있기 때문에 이를 보완할 수 있다. 순환총조사가 제공해줄 것에 대한 명확한 서술 없이 사용자가 순환총조사의 이점들을 충분히 평가하기는 어렵다. 사용자들이 순환총조사를 총조사 방법으로 이해하고 지지하기 위해서는 순환총조사의 이점에 대한 포괄적인 설명이 필요하다.

3. 연속시행

연속실행(continuous operation)은 조사의 효율성 증가와 품질향상에 도움이 된다. 순환총조사는 연중 지속적으로 시행되기 때문에 현행 총조사와 같은 주기 조사 환경에서는 달성하기 어려운 문제, 예를 들면 현장조사인력의 고용 안정과 같은 문제를 해결할 수 있다. 현장조사인력의 고용이 안정되면 이들의 경험이 축적되어 조사의 효율성이 높아지고 조사품질이 향상될 것이다. 또한 지속적인 자료의 생산은 통계 이용자의 요구를 수시로 수용하여 지속적인 방법론의 개선을 이룰 수 있다.

하지만 연속시행 조사는 새로운 위험성이 제기될 수 있다. 예를 들어 선거와 같이 지역과 연관된 일이 있을 경우 이러한 외부 요인이 총조사 과정에 개입할 위험이 있으며, 총조사 결과에 영향을 미칠 수 있다.

4. 총조사의 브랜드 가치 하락

순환총조사와 전통적인 총조사의 주요한 차이 중의 하나는 집중적인 전국 선전 캠페인 효과가 사라진다는 것이다. 전통적인 총조사의 브랜드 가치는 매우 높으며, 이것은 총조사의 큰 이점이다. 순환총조사는 비록 국가사업의 한 부분이긴 하지만 현행 총조사와 같은 전국적인 대규모 사업은 아니기 때문에 총조사의 브랜드 가치는 떨어질 것이다. 브랜드 가치 하락은 공공 수요의 감소를 유발할 수 있다. 비록 순환총조사의 부정적 전국 매체보도나 반-총조사 그룹과 같은 문제가 발생하지 않는다 하더라도 순환총조사의 브랜드

가치 하락은 유념해야 할 점이다.

5. 포함 범위

순환총조사는 공공 수요의 감소 이외에도 포함범위(coverage) 문제를 야기할 수 있다. 예를 들어, 인구의 흐름은 유동적이기 때문에 총조사 주기 동안 한 번 이상 포함된 가구의 비율이 상당할 수 있고, 반대로 주기 동안 표본이 되는 지역에 한 번도 거주한 적이 없는 가구는 포함되지 않는다는 문제도 있다. 또한 항상 표본이 되는 지역 밖에 산다고 되어있으나 다른 장소에서 보통 거주자처럼 측정되는 이중 거주자 문제도 있다. 이러한 추가적인 포함범위 문제들에 대한 검토가 필요하다.

6. 법령 제정

현장 조사의 빈도가 늘어나므로 순환총조사 설계를 적용하기 위해서는 새로운 법적 기반이 필요하다. 현재 총조사는 총조사가 실시되는 기간, 지역, 주제, 범위 등이 법률 혹은 시행령 등에 명시되어 있다. 따라서 순환총조사로 조사 방법을 변경하면 관련 법률을 개정하고 새로 제정해야 한다.

7. 비용

순환총조사는 전통적인 센서스보다 더 많은 비용이 들 것으로 보인다. 어느 정도의 비용절약은 가능하지만 순환총조사 수행에서 처음에 비용이 많이 들게 될 것이다. 방법론, 진행과정, 관리체제 구축, 관련 인력 정비 등 초기 비용은 몇 년의 운영 후에야 그 효과성을 나타나게 될 것이다. 이러한 초기 비용으로 말미암아 연속성에 의한 비용 절약은 아주 작은 부분일 수도 있다. 이런 이유에서, 주기적인 총조사보다 순환총조사는 총비용 면에서 더 많은 비용을 필요로 할 가능성이 많다.

제4절. 우리나라 실정에 맞는 순환총조사 방법 검토

순환총조사를 우리나라 도입하는 방법은 두 가지가 있다. 프랑스의 경우와 같이 순환총조사를 전면적으로 도입하는 방법과, 미국지역사회조사와 같이 총조사의 표본조사에 순환총조사를 도입하는 방법이 있다. 두 방법 모두 아래 사항을 공통적으로 검토하여야 한다.

1. 예산 및 인력

가. 예산

순환총조사를 도입하면 조사예산이 5년으로 분산되나 5년 누적 총예산은 증가할 것이다. 통계청에서는 총조사 실시 비용 및 지방자치단체 보조금 예산이 필요하고, 지방자치단체는 총조사를 연중 실시하는 예산이 필요하다.

나. 인력

총조사를 연중 진행할 인력을 확보하여야 한다. 통계청에서는 총조사를 진행할 중앙 관리자가 필요하고, 지방자치단체는 총조사 실시 관리자 및 현장 조사 인력이 필요하다.

2. 순환총조사 방법론 개발

가. 프레임 구축

순환총조사를 실시하기 위해서는 양질의 프레임을 구축하여야 한다. 포함 범위가 100%에 근접한 프레임을 구축하는 것은 매우 중요한 작업이다. 이때 결정해야 하는 사항으로는 다음과 같은 것이 있다.

- 단위 기준 : 가구 혹은 거처
- 표본추출단위 : 조사구 혹은 다른 기준, 조사구 크기 문제
- GIS 활용 가능 여부

나. 프레임 균등 분할

순환총조사를 실시하기 위해서는 프레임을 균등하게 분할하는 기법을 개발하여야 한다. 이때 결정해야 하는 사항은 다음과 같다.

- 전국을 균등하게 5등분 하는 기법 개발
- 이때 활용 가능한 인구통계적 정보 결정 : 성별, 나이, 학력 등
- 균등 분할하는 최저 기준 : 조사구, 읍면동, 혹은 시군

다. 프레임 변동의 체계적인 갱신

순환총조사가 실시되는 5년 조사주기 동안 이사, 출생, 사망 등으로 프레임은 계속적으로 변동한다. 순환총조사가 성공하려면 프레임 변동이 지속적으로 순환총조사에 반영되어야 한다.

- 이사, 거주지 이동 등 거주지 이동
- 출생, 사망, 결혼, 분가 등으로 발생하는 인구이동

이를 위해서는 프레임 변동이 기록되는 지방자치단체의 협조가 필수적이다. 주민등록 변동, 이사 등으로 인한 거주지 변동 등이 체계적으로 지방자치단체로부터 통계청으로 통보되어야 한다.

라. 모수추정 기법 개발

순환총조사에서는 일년 추정치뿐 아니라 3년 혹은 5년 추정치가 동시에 제공된다. 통계청에서는 공표하는 추정치에 대한 구체적인 방법론을 개발해야 한다.

- 연차 데이터의 누적 방법 연구 : 이동 평균법, 가중 방법 등
- 5년 누적 데이터를 이용한 인구수 및 가구수 추정 기법
- 공표하고자 하는 결과의 지리적, 인구통계적 범위 결정
- 가구 및 인구의 이동상황에 대한 모형화 기법 연구

3. 지방자치단체의 협조 필요

순환총조사가 성공적으로 실시되려면 통계청과 지방자치단체와의 파트너십이 무엇보다 중요하다. 현행 총조사와는 달리 순환총조사는 많은 부분을 지방자치단체로부터 협조를 받아야 하기 때문이다. 지방자치단체와의 파트너십이 필요한 구체적인 부분은 다음과 같다.

- 프레임 갱신에 필요한 자료 협조
- 지역수준에서 행정 자료의 적절한 협조 필요
- 순환총조사 실시에 필요한 조사자, 감독관의 역할 분담

4. 순환총조사에 대한 홍보 및 법령 제정

가. 지속적인 홍보

순환총조사를 실시하려면 일반국민이 순환총조사를 받아들일 수 있도록 사전홍보가 필요하다. 대부분은 현행 총조사에 익숙해 있으므로 새로운 조사가 도입되기 위해서는 이를 알려서 거부감 없이 받아들일도록 지속적인 홍보를 하여야 한다. 홍보는 정부, 지방자치단체, 학계, 관련 단체들을 세분하여 이들의 이해를 높이는 것이 효과적일 것이다.

- 정부 정책결정자의 이해 제고
- 지방자치단체, 학계, 관련단체의 이해 제고
- 일반 국민의 이해 제고

나. 법령 제정

순환총조사를 실시하려면 순환총조사를 실시 할 수 있는 법적 토대를 구축해야 한다. 관련 법령을 제정하여 법적 테두리 안에서 총조사를 실시하여야 한다.

제5절. 소결

순환총조사는 현행 총조사의 대안으로서 장점과 단점을 동시에 가지고 있다. 장점은 최신 자료를 시의성 있게 생산하고, 일 년에 집중되는 총조사 예산을 5년으로 분산하여 예산 확보의 어려움을 분산할 수 있는 점이다. 그리고 총조사를 지속적으로 실시함으로써 방법론 개선 및 현장 인력의 경험 축적 등 총조사의 효율성과 품질 수준을 높일 수 있다. 반면, 총 예산은 도리어 증가할 가능성이 높고, 지방자치단체로부터 긴밀한 협조가 필수적이며, 새로운 방법론에 대한 개발하고 관계 법령을 제정해야 하는 어려움이 있다. 또한 정부의 정책 결정권자 및, 관련 단체, 학계에 이해를 높여야 하고 일반 국민을 대상으로도 홍보를 통하여 이해를 높여야 한다. 따라서 순환총조사를 적용하기 위해서는 이러한 장단점을 종합적으로 검토하여 체계적인 이행 계획을 세운 뒤 실행에 옮겨야 할 것이다.

제8장. 결론 및 제언

제1절. 인구주택총조사의 활용성

1. 설문조사

이 연구에서는 먼저 인구주택총조사가 어떻게 이용되는지를 알아보기 위해 설문조사를 실시하였다. 조사대상은 공공부문인 중앙부처와 지방자치단체의 공무원과 민간부문인 연구 및 교육기관의 연구자로 하였다. 공공부문에서는 7급 이상 3급 이하 직급의 공무원을 각각의 홈페이지 조직도에서 찾아 표집틀을 만들었다. 여기서 부처별·지자체별로 30명 씩 무작위 추출하여 중앙부처에서 1,240명, 지방자치단체에서 539명을 표본으로 하였다. 민간에서는 관련 학회 및 연구소 명단과 통계청에서 확보한 이용자 명단을 이용하여 90명의 표본을 만들었다. 조사에서는 온라인 조사와 오프라인 조사를 병행하여 총 응답자수는 중앙부처가 177명으로 응답률 14.3%, 지방자치단체가 173명으로 32.1%, 민간인 18명으로 20.0%였다.

조사결과, 지방자치단체가 중앙부처보다는 인구주택총조사를 조금 더 적극적으로 이용하고 있는 것으로 나타났다. 또한, 지방자치단체의 총조사에 대한 요구도 중앙부처에 비해서 더욱 적극적이었다. 요구의 주요 사항은 이용편의, 소지역 통계, 세부 통계, 시의성 있는 통계 등이었다.

이런 측면에서, 먼저 인구주택총조사의 활용성을 높이기 위해서는 지방자치단체에서 필요로 하는 통계에 대해 좀 더 심각하게 고민해 볼 필요가 있다. 기실, 중앙부처는 행정에서 주민등록을 이용하기 때문에 인구주택총조사 통계에 대한 수요가 상대적으로 적은 편이다. 반면, 지방자치단체에서는 지역의 세부사항을 파악하기 위한 통계가 점점 더 필요해지고 있다. 이런 수요에 통계청이 적극적으로 대처하기 위해서는 인구주택총조사에서 지방자치단체의 통계수요를 파악하는 데 좀 더 많은 노력을 기울일 필요가 있다.

한 가지 부연할 것은, 아직 통계에 대해 그다지 많은 경험을 갖지 못한

지방자치단체의 경우 자신들이 어떤 통계를 필요로 하는지 잘 알지 못하는 경우가 있을 수 있다. 이런 경우에 대비하여 지방자치단체의 통계교육을 강화하고, 통계와 관련된 문제를 함께 상의하며, 그들의 통계문제에 적극적으로 상담해주는 협력체계를 정립하는 것이 바람직하다. 이에 본 연구는 통계청의 통계서비스에 지방자치단체나 중앙부처에서 필요한 통계에 관해 상담할 수 있는 조급은 전문화된 상담창구를 제안한다. 상담창구를 통한 전문화는 통계서비스를 차별화하고, 공공부문 통계서비스의 수준을 질적으로 높이고, 더 나아가 정책에 필요한 통계수요를 파악하는 데 도움을 줄 것이다.

조사와 관련하여 마지막으로, 이 연구는 인구주택총조사 이용자들의 통계요구를 파악하고, 그에 부응하여 새로운 센서스 방법들을 모색하기 위한 것이다. 이런 측면에서, 이 연구는 통계 생산자의 입장보다는 수요자 중심에서 인구주택총조사의 미래를 생각한다는 데 최소한의 의의가 있다. 앞으로 인구주택총조사를 포함한 많은 국가통계는 수요자 입장에서 평가하는 통계의 효용가치가 그 통계의 지속적 생산 가능성을 담보해 주게 될 것이기 때문이다. 따라서 앞으로 통계수요를 파악하는 좀 더 세련된 조사방법을 개발할 필요가 있다는 점도 함께 제시하고자 한다.

2. 활용 증진 방안

설문조사결과, 특히 인구주택총조사에 대한 요구 사항들을 바탕으로, 이 연구는 몇 가지 제안을 하고자 한다. 우리나라에서는 행정에서 주민등록자료를 이용하기 때문에 인구주택총조사의 이용확대를 기대하기는 어려울 것으로 보였다. 그렇지만 설문조사는 인구주택총조사에 대한 지방자치단체의 고유한 수요가 있다는 것을 보여주었다. 지역사회의 특성은 주민등록으로 파악할 수 없을뿐더러, 인구주택총조사와 같은 대규모 조사가 아니면 파악하기 힘들기 때문이다. 또한 행정수요의 증가는 행정학이나 사회학 분야에서 지역사회 연구의 증가로 이어지면 이에 따라 학술연구에서도 총조사에 자료에 대한 수요가 늘고 있다. 이러한 소지역 통계의 수요 가운데 특히 총조사의 표본조사에 대한 소지역별 통계에 대한 요구가 가장 많다.

이에 대한 해결 방안으로 이 연구는 공표집계구 설정과 이를 포함한 통계구역체계의 정립을 제안한다. 이를 위해 외국의 통계구역 사례를 살펴보고, 이를 바탕으로 현재 진행되고 있는 우리나라의 통계구역설정을 살펴보았다. 그 가운데 특히 소지역통계 제공에 중심 역할을 하는 집계구를 중점적으로 살펴보았다. 집계구는 통계를 가능한 가장 작은 지역 수준까지 생산함으로써 지방자치단체의 정책에 직접 활용될 수 있는 것이다. 이러한 통계는 단지 행정뿐만 아니라 민간기업의 마케팅, 입지선정, 표본조사의 표집틀 등으로 다양한 활용성을 잠재하고 있다. 또한 국가 중심이었던 학술적 연구가 지역별로 비교연구로 확대될 수 있는 가능성을 제공한다.

한편 집계구 및 기타 통계구역의 기반이 되는 기초단위구는 행정의 또 다른 수요가 있어 주목할 만하다. 도농복합의 성격으로 행정구역을 정의하는데 어려움이 있는 지역들이 많이 나타나고 있다. 이런 지역에서 통계에 바탕을 두고 행정구역을 나눈다면 좀 더 현실적이며 효율적인 행정의 근간을 마련할 수 있다. 즉, 도농복합지역에서 인구주택총조사는 행정구역 지정 및 획정의 근거로 활용될 수 있다.

이러한 작업들은 모두 GIS의 기술적 지원이 또 다른 한 축을 지탱하고 있다. 다시 말해, 전국을 작은 지역 단위로 나누어 그것을 통계의 목적으로 재결합하는 작업은 GIS의 지원이 없다면 구현하기 매우 어려운 것이다. GIS의 이용은 다른 한편, 통계 자료와 지도의 결합을 의미하기도 한다. 이제 지도와 통계 자료는 하나의 DB로 연계·구축되는 것이다. GIS 기술을 기반으로 지역들을 통계적 동질성과 공간적 근접성으로 함께 묶으면, 통계와 지역이 함께 쌍을 이루는 소구역들이 만들어진다. 이러한 소구역들은 지도에 표시하기도 쉽고, 사람들이 이해하기 쉽기 때문에 인구주택총조사의 이용을 촉진시키고 확대하는 데 많은 역할을 할 것이다.

또 하나의 제안은 원자료 이용을 활성화하는 것이다. 통계청의 통계 서비스가 줄곧 발달해왔지만 그 가운데 원자료에 관한 부분이 상대적으로 뒤떨어져 있다. 원자료 서비스는 MDSS와 CD와 같은 매체를 통해 어느 정도 이루어지고 있으며, 비밀보호가 필요한 자료에 대해서는 on-site 서비스도 작은 규모로나마 이루어지고 있다. 그렇지만 아직은 이에 대해 뚜렷한 정책의

방향이 정립 돼있지 않다. 여기서는 자료 형태에 따라 Public Use Microdata와 Confidential data를 구분하여, 서비스를 차별화하는 것을 권한다. 거의 모든 선진국들이 이와 같은 방식으로 원자료를 서비스하고 있다. 또한 PUM은 무료로, Confidential data에 대해서는 이용 절차와 규정을 까다롭게 적용하며 비용을 책정하는 것이 전반적 경향이다. 이는 통계 이용자가 전문화되어감에 따라 서비스도 전문화되어 가는 것을 의미하기도 한다.

한편, 원자료 서비스는 정보화 시대의 국가 간 경쟁에서 국민들의 정보역량을 높이는 데 필요한 기초 자원을 국가 제공하는 GOV2.0에 적극적 참여를 의미한다. 이제 통계청은 국가의 가장 기초가 되는 정보를 생산하는 작은 역할에서 벗어나, 이의 유통 및 활용을 장려하여, 국가 경쟁의 초석을 제공하는 큰 역할로 자신의 모습을 바꾸어 나아가야 할 것이다.

제2절. 행정자료 이용

행정자료 이용 현재 모든 국가가 관심을 기울이는 새로운 센서스 방식이다. 물론 이의 최종 목표는 행정자료로 조사를 대체하는 것이다. 그렇지만 북구유럽을 제외하고는 아직 행정자료 기반 센서스를 본격적으로 실시하는 나라는 거의 없다. 북구유럽에서도 덴마크와 핀란드는 이미 오래 전부터 행정자료로 모든 센서스 통계를 대체했지만, 스웨덴은 지속적으로 행정자료로 센서스를 대체해 나아가고 있으며, 노르웨이는 2011년에 완전 대체를 목표 현재 거처주소(dwelling household project)를 정비하고 있다.

2000년대에 들어 우리나라에서도 행정자료로 센서스를 대체하자는 의견이 빠르게 전파되고 있다. 여기에는 1962년 제도화된 우리의 주민등록제도가 이제 상당히 안정되어 행정의 전 분야에서 사용되는 것이 가장 큰 이유다. 또한 1990년대 말과 2000년대 초에 부동산 관련 행정자료, 특히 건축물대장 자료의 정비도 행정자료로 센서스를 대체할 수 있다는 의견에 큰 몫을 차지한다. 이런 의견들이 얼마나 현실적으로 가능하지를 알아보기 위해 인구주택총조사자료와 주민등록자료를 전국, 시도, 시군구, 읍면동 수준에서 비교해 보고, 2개의 지역에서 2004년에 실시한 인구주택총조사 시범조사 자료와 같은

시기의 주민등록자료를 개인 수준에서 비교해 보았다. 또한 건축물과 관련 있는 행정자료들과 이들 사이의 연계 가능성에 관한 기존 연구들을 발굴하여 정리했다.

그 결과 지금 당장은 행정자료로 인구주택총조사의 조사 항목에 대한 통계를 대체할 수 없다는 사실을 알 수 있었다. 일단 인구주택총조사와 주민등록자료의 인구 및 가구-세대 차이가 컸으며, 행정구역별 집계수준이 작아질수록 총조사와 등록 자료의 차이는 커지고 편차도 커졌다. 인구의 경우 전국 수준에서는 전체적으로 4% 정도의 차이가 났지만, 읍면동 수준에서는 전체 3,572개 동 가운데 1,239개(34.7%)가 10% 이상 차이가 났으며, 30% 이상 차이가 나는 읍면동이 40개 가량 있었다. 또한 이상값을 제외하더라도 가장 큰 차이는 60% 이상이었다. 가구-세대의 차이는 이보다 더 커서 전국 수준에서는 10% 정도의 차이가 났고, 읍면동 수준에서는 90% 이상 차이가 나는 지역도 있었다. 이와 같은 현상은 연령별 차이에서도 마찬가지였다. 연령별 분석에서 가장 두드러진 결과는 15-24세 사이의 연령층에서 주민등록과 총조사 자료 사이의 차이가 큰 것이다. 이는 젊은 시절에 대학이나 직장으로 거주지를 옮기는 경우가 많은데 이때 주민등록의 주소는 바꾸지 않기 때문에 나타난 것으로 이해할 수 있다.

이 연구에서 나타난 주민등록과 인구주택총조사의 차이는 앞으로 주민등록으로 총조사 인구를 대체하기 위해서는 이러한 차이를 보정할 수 있는 모형의 개발이 필요하다는 것을 제시한다. 이를 위해서는 먼저 인구에서는 인구주택총조사와 같은 시점의 주민등록 자료를 사용하고, 총조사와 같은 집계 방법으로 집계하여 비교해 보아야 할 것이다. 더 나아가 이러한 자료를 시계열로 비교해 보아야 할 것이다. 다음은 이러한 비교 결과를 토대로 보정 모형을 개발해야 한다. 이때 연령별 특성과 지역의 특성은 모형에 중요한 역할을 할 것으로 보인다.

보정 모형 개발의 최종 단계에서는 개인 수준에서 두 자료를 비교하는 과정을 거쳐야 한다. 이러한 작업을 통해 개인 수준에서 발생하는 주민등록과 실제 거주 차이에 대한 좀 더 확실한 이해가 있어야 신뢰할 수 있는 보정 모형을 만들 수 있을 것이다. 더욱이 이 연구에서 다른 자료가 읍면동 수준

이 가장 작은 집계 수준이어서 집계구와 같이 이보다 작은 지역의 집계를 내기 위해서는 좀 더 세밀한 모형이 필요하기 때문이기도 하다. 이를 위해서는 2008년 시험조사서부터 본 조사, 더 나아가서는 본 조사 이후에도 여러 차례에 걸쳐 주요 지역들, 특히 지속적으로 두 자료의 차이가 많이 나는 지역부터 표본을 추출하여 개인 수준에서 직접 주민등록과 조사 자료를 비교하여야 할 필요가 있다. 이때 적어도 가구주의 이름, 가구원의 생년월일, 거처의 주소를 함께 조사하여 주민등록과 직접 비교가 가능하도록 해야 할 것이다. 또한 가능하면 가구원의 이름을 두 자리만이라도 받을 수 있으면 더욱 좋을 것이다.

건축물 관련 행정자료는 인구주택총조사에서 가구와 주택에 관한 사항을 대체하는 데 이용될 수 있다. 그렇지만 현재는 건축물 관련 행정자료 들 사이의 문제, 자료 자체의 정확성, 서로 다른 분류 체계 등으로 사실상 이용이 불가능하다. 가장 시급히 개선할 문제는 현재 사용하는 주소인 필지번호(지번)다. 현재 필지 중심의 주소 체계로는 건물에 고유한 번호를 부여하기 어렵다는 점이다. 이는 단지 통계의 문제뿐만 아니라 이미 도로 중심으로 건물에 번호를 부여하는 새주소가 만들어지고 있다. 또한 새주소법이 국회를 통과하면 공식적으로 행정에서도 새주소를 사용하게 되어 건물의 고유한 번호를 사용하기는 점점 나아질 것이다.

그렇지만 새주소로도 행정자료에서 총조사의 가구정보를 취득하는 데는 한계가 있다. 건물 내부의 거처별로 고유한 번호가 있어야 이를 바탕으로 같은 거처에 사는 사람들을 인식할 수 있고, 그래야만 이들을 하나의 가구로 정의할 수 있기 때문이다. 스웨덴과 노르웨이가 행정자료 기반 센서스로 나아가는 데 어려움을 겪고 있는 데는 바로 거처별로 고유한 주소를 갖지 못한 것인 큰 원인이었다. 이런 이유에서 노르웨이는 거처별로 주소를 부여하는 거처주소프로젝트를 실시하여 진행 중에 있으며, 이를 바탕으로 2011년에는 완전히 행정자료 기반의 센서스가 될 것을 기대하고 있다.

새주소가 정착되어 건축물 관련 행정자료를 서로 연계할 수 있다고 하더라도, 만약 개별 자료가 부정확하다든지 누락되어있다면 이 또한 행정자료로서 사용가치가 매우 낮아진다. 또한 건축물 관련 항목의 분류체계나 항목 자

체의 의미가 다르면 이 또한 사용하기가 어렵다. 따라서 건축물 관련 행정자료로 총조사를 대체하기 위해서는 개별 행정기관에서 정확히 자료를 입력하고, 적정한 주기로 갱신해야 하며, 더 나아가서는 같은 분류체계와 같은 용어를 사용해야 한다. 그렇지만 이 모두가 통계청이 할 수 있는 일이 아니기 때문에 모든 부처가 공히 사용할 표준안을 만들고 이를 사용하기까지는 그다지 쉽지 않을 것이다.

이런 어려움에도 불구하고, 가구와 주택부분의 인구주택총조사 항목 가운데 건축물 관련 행정자료와 통계청의 누적된 인구주택총조사자료를 함께 이용하면 인구주택총조사의 응답부담을 줄일 수 있는 부분이 있다. 지난 2005년 총조사에서 가구와 주택에 관한 항목은 통계청의 DB에 들어가 있다. 2005년 이후 신축하거나 개축한 경우와 변경하거나 부순 경우 이외에는 이들은 변하지 않았다. 이러한 항목은 가구에 관한 사항 가운데 사용 방수, 주거 시설 형태, 거주층, 난방시설이며, 주택에 관한 사항은 거처의 종류 및 건물층수, 연건평, 대지면적, 총방수, 건축년도, 편익시설수가 있다. 먼저 건축물대장과 건축허가신청서 등을 이용하여 건축물 신축, 개축, 변경, 붕괴 등을 점검한다. 변동이 없는 건축물에 대해서는 과거의 자료를 그대로 사용하고, 변동 사항이 있을 경우도 조사에서는 변동된 사항을 확인하여 수정하면 응답자 부담을 줄일 수 있을 뿐만 아니라, 조사비용도 줄일 수 있을 것이다. 물론 이러한 과정은 시험조사를 통해 실제로 어떤 문제가 있는지를 파악하고, 특히 조사원 교육에서 발생할 문제를 미리 점검하는 것이 필요할 것이다.

행정자료 이용에서 최종의 목표는 행정자료로 모든 센서스 통계를 생산하는 것이다. 이를 위해서는 건축물 관련 행정자료에서도 논의하였듯이 많은 작업이 필요하며, 더욱이 그러한 작업은 통계청이 혼자서 할 수 있는 것이 아니다. 따라서 행정자료 기반의 센서스를 추진하는 자세는 길게 내다보고, 관련 행정부처에 협조를 구하고, 더 나아가 국가적 차원에서 협조체계를 구축하고, 행정절차에 통계가 스며들도록 하는 것이다.

앞으로 인구주택총조사에서 행정자료를 활용하기 위해서, 더 나악 행정자료 기반의 센서스로 옮겨 갈 수 있기 위해 지금 해야 할 사안들은 다음과

같다.

- 통계청이 확보한 기존 자료를 최대한 이용할 수 있는 체계구축
- 보완적으로 활용 가능한 행정자료의 발굴 및 이용체계 구축
- 새주소체계 활용 및 거처별 번호 부여 작업
- 수치지도와 거처별 번호 연계
- 조사자료에 거처별 번호 입력
- 조사자료와 행정자료와 비교 연구

제3절. 순환(총)조사

순환(총)조사에 대해서는 프랑스의 순환총조사와 미국의 순환조사를 살펴 보았다. 아직 순환(총)조사가 시행 중에 있으며, 총체적인 결과가 나오지 않아 순환(총)조사방법에 대해 평가하기는 이르다. 현재까지 두 국가에서 시행해 온 것을 토대로 우리나라의 인구주택총조사에 어떤 시사점을 주는지 살펴 보았다.

먼저 순환조사를 프랑스와 같이 전수조사에 적용할 것인지 아니면 미국과 같이 표본조사에 적용할 것인지를 검토해야 한다. 결론부터 이야기 하자면, 우리 실정에 전수조사보다는 표본조사가 더 적합하다. 왜냐하면, 현재 인구주택총조사를 시행하는 데 문제가 있는 것은 아니기 때문에 5년에 한번 하는 전수조사를 굳이 바꿀 필요는 없을 것이다. 다만 5년에 한번 결과가 발표되어 통계의 시의성이 떨어지는 문제는 고려해 볼 만하다. 특히 지방자치단체가 총조사 통계에 대해 더 적극적이며, 그 가운데도 표본조사로 만들어지는 소지역 통계에 대한 수요가 크다. 이런 측면에서 표본조사를 순환조사로 하면 매년 소지역 통계를 생산할 수 있을 것이다. 이때 소지역의 수준이 문제인데, 매년 생산하는 통계는 시군구 수준, 5년을 합한 통계는 읍면동 수준 까지 생산이 가능할 것이다.

지금 10% 표본조사를 그대로 시행한다고 가정하고 먼저 읍면동 수준에서 추출한다면, 전국의 읍면동이 3,584개니 표본으로 뽑히는 읍면동은 약 360개가 된다. 5년 동안 매달 조사를 한다면, 60개월의 조사가 되고, 한 달에 6개

의 읍면동을 조사하게 된다. 가구로 보면 10% 표본이 150만 가구니, 한 달에는 25,000 가구를 조사하는 셈이다. 이를 종합하면, 평균적으로 한 달에 한 동에서 420 가구씩, 6개 읍면동을 조사하면 현재와 같은 10% 표본을 조사할 수 있다.

그 다음 문제는 시군구 수준에서 대표성 확보를 몇 년 주기로 하느냐는 것이다. 복잡한 경우는 인구수와 행정구역 수가 많은 경우다. 예를 들면 서울에는 25개 구가, 경기도에는 31개 시군구가 있다. 이러한 시도에서 시군구 전체를 대표하는 자료를 매년 생산할 것인지, 아니면 매년 일부 시군구만 대표성 있는 자료를 생산하고 5년을 합쳐 시군구를 대표할 자료를 생산할지, 아니면 2-3년의 주기로 할지 등을 정해야 한다. 현재로는 이에 대해 어느 것이 좋다고 정할 수는 없으나, 표본조사를 순환조사로 할 경우에는 먼저 결정해야 할 사안이다.

다음은 방법론의 개발이다. 특히 표본틀(sampling frame)이 계속 변하기 때문에 이에 대한 대책이 있어야 한다. 이에 대해서는 현재와 같이 주민등록의 인구변동을 가져다 인구를 추정하는 수밖에 없다. 물론 이를 위해서는 행정부와 지방자치단체에서 정기적으로 주민등록의 변동 사항을 통계청에 통보해 주어야 한다. 따라서 이 부분에서도 행정자료의 이용이 필수적인 사안이 될 것이다.

그렇지만, 현재 인구이동통계를 내는 경우와는 고려해야 할 사항이 다르다. 인구이동통계는 발생한 현상, 즉 과거의 것을 통계에 반영하는 것이다. 반면, 표본설계는 미래의 것을 준비하는 작업이다. 적어도 조사시점의 12달 전에는 1년 조사의 표본설계가 완료되어있어야 하며, 이상적인 경우는 60개월 전체에 대한 표본설계가 있어야 할 것이다. 이렇게 미래의 지역별 인구를 추정하기 위해서는 지역에 대한 예측 가능한 사건, 예를 들면 재개발인 택지 개발 등에 대한 고려를 해야 할 것이다. 또한 예측이 불가능 사건, 예를 들면 재해나 천재지변 등이 발생하여 인구가 감소(증가)하는 경우 어떻게 대응할 것인지에 대한 절차 및 방법에 대한 사전의 준비가 있어야 할 것이다.

이외에도 심각한 방법론의 개발이 요구되는 사안들이 많다. 특히 현재 표본조사에서 사용하는 표집단위(sampling unit)인 조사구로 대표성을 확보할

수 있는지는 의문이다. 조사구들 사이의 표본 수 차이가 너무 크기 때문이다. 따라서 표본조사를 순환조사로 할 경우 현재의 조사구보다는 좀 더 일관된 크기의 새로운 표집단위의 개발해야 한다. 이는 지금까지 통계청이 해온 모든 표본조사의 근간을 바꾸는 작업이다. 경제활동인구조사나 가계조사와 같은 통계청 조사뿐만 아니라 통계청으로부터 표본의 협조를 받는 다른 부처의 표본조사에도 큰 변화가 따른다는 점도 고려해야 할 것이다.

이와 더불어 모수추정에 대한 새로운 방법론 개발이 필요하다. 연차 자료의 누적방법에 대해 이동 평균법과 가중 방법을 개발해야 하며, 5년 누적 데이터를 이용하여 인구수 및 가구수를 추정하는 방법을 개발하여야 한다. 또한 가구와 인구가 이동하는 상황에서 모수를 추정하는 모형을 개발해야 한다. 또한 공표하고자 하는 결과의 지리적 및 인구통계적 범위를 결정해야 한다.

다음으로 고려할 문제는 상시조사체제가 되기 때문에 이 조사에만 전념하는 전문요원이 필요하다는 것이다. 예를 들면 하루에 10 가구씩 방문조사를 한다고 하면 한 달에 한 조사원이 20일을 조사에 투입한다고 보았을 때 200 가구를 방문하여 조사할 수 있다. 이럴 경우 125명의 표본조사 전담 조사원이 필요하다. 이것은 인력에 대한 예산과 함께 지금 한창 문제가 되는 비정규직 고용과도 직접 관계가 있다. 이런 점들을 감안해보면, 전체적인 예산은 현재의 표본조사에 비해 많이 들 것이며, 특히 초기에 방법론 개발과 조사체계를 안정되게 구축하는 데 비용이 많이 들 것이다. 더불어 여러 차례의 시험조사와 본조사를 비교연구하고, 이러한 과정을 여러 차례 거쳐 적절한 방법론을 개발하는 장기간에 걸친 사업으로 진행하여야 유용한 통계가 생산될 수 있다는 점도 염두에 두어야 할 것이다.

간략하게 2015년 인구주택총조사의 표본조사를 순환조사로 대체한다고 할 때 이에 대비해 해야 할 최소의 사항들을 열거하면 다음과 같다.

- 2010년 표본조사와 비교를 위한 2009년, 2010년, 2011년의 조사
- 2009년 조사를 위한 여러 차례의 2008년 시험조사 및 평가
- 조사구보다 작으며, 어느 정도 일정한 규모의 표집단위 설정 및 구축

제4절. 제언

이 연구의 주된 관심은 우리나라의 센서스가 앞으로 어느 방향으로 가야 할 지를 탐색하는 것이다. 현재 우리가 직면한 상황은 행정자료 기반의 센서스와 순환조사 가운데 어느 쪽을 선택할 것인지를 결정하는 것이다. 물론 이 연구 하나로 앞으로의 방향을 결정하는 것은 매우 위험하다. 따라서 이 연구에서는 가능하면 다각적이며 정확한 근거를 가진 제안을 하여 의사결정자들이 의사결정에 도움을 주도록 하는 것이 최선일 것이다.

먼저 행정자료 기반의 센서스는 많은 국가들이 현재의 센서스를 대체할 미래의 센서스 방법으로 생각하고 있다. 현재는 북구유럽의 몇 국가들이 실시하는 정도에 불과하지만 다른 나라들도 행정자료 기반의 센서스를 향해 여러 가지 방법들을 시도해 보고 있다. 그 가운데 행정자료를 조사에서 보완적으로 사용하는 경우가 가장 일반적이다.

한편 순환조사는 단지 두 나라가 시행하고 있을 뿐이어서 행정자료 기반의 센서스에 비해 그다지 보편적인 것은 아니다. 행정자료 기반의 센서스는 이 방법을 20년 이상 시행한 국가들이 있기 때문에 어느 정도 검증도 되었고, 방법론을 배울 수도 있다. 이에 반해 순환조사는 아직 최종의 결과가 나오지 않은 상태여서 조사에 대한 검증이 이루어지지 않았고 독자적으로 개발해야 할 방법론적 과제가 많다.

이런 국제적 경향을 보면, 우리가 왜 이 시점에서 센서스의 향방을 생각해야 하는지부터 살펴볼 필요가 있다. 먼저 우리나라에서 센서스 시행에 아직 커다란 문제가 생긴 것은 아니다. 독일이나 네덜란드처럼 응답거부가 많은 것도 아니고, 호응도 그다지 나쁘지 않다. 특히 지난 총조사에서 홍보는 조금은 축제 분위기를 만들었던 것 같다. 이런 측면에서 일반 국민들이 생각하는 센서스는 현재로는 그다지 큰 문제는 아닌 것으로 보인다.

반면 총조사의 비용증가와 결과의 유용성에 대한 비판은 조금은 심각하게 다가오고 있는 것으로 보인다. 비용증가는 인구증가로 인한 조사비용 증가, 노인가구와 단독가구의 증가로 인한 조사비용 증가, 지방자치단체의 통계기능 약화에 따른 비용 증가 등 여러 원인이 있지만 그러한 원인은 어쩔 수

없는 것이고 그 때문에 인구주택총조사를 비판하는 것은 아니다. 여기에는 ‘인구주택총조사에 그 많은 비용을 들여서 어디에 쓰는지’로 대표될 수 있는 비판, 즉 비용증가에 따른 유용성에 대한 회의가 있기 때문이다. 이런 측면에서, 우리가 인구주택총조사의 앞날을 생각할 때는 비용절감의 측면과 유용성 확대의 두 면을 함께 고려해야 할 것이다.

행정자료 기반의 센서스를 생각해 보면, 이 방법이 완전히 정착될 경우 장점이 단점에 비해 훨씬 우월하다. 다만, 우리나라가 과연 이러한 통계생산의 체계를 갖출 수 있을지에 대해 확신이 서질 않는다. 왜냐하면, 이러한 체계를 갖추고 정착시키는 데는 통계정보보다 훨씬 넓은 범위, 때로는 많은 행정부처와 지방자치단체, 더 나아가 국가의 모든 기관과 국민들 적극적인 협조내지는 동의가 필요할 수도 있기 때문이다. 더욱이 이런 체제가 제대로 정착되는 데 얼마나 오랜 시간이 걸릴지 아무도 예측할 수 없는 것도 문제다.

이에 반해 순환조사는 통계청이 혼자서 수행하기에 상대적으로 수월하다. 이런 측면에서 당장 순환조사에 대한 욕심이 생기는 것도 어쩔 수 없는 인지상정이다. 특히 요즈음처럼 인구주택총조사에 대해 시의성 있는 통계, 소지역 통계의 요구가 커진 상황에서 순환조사를 시행하면 이러한 요구에 부응할 수 있을 것도 같으며, 더욱이 통계청이 독자적으로 통계 수요에 적극적으로 대처하는 모습을 보여 줄 수 있어 매력적이다.

이렇게 보면, 행정자료 기반의 센서스인가 아니면 순환조사인가는 통계수요에 대한 장기적 대응인가 단기적 대응인가를 선택하는 문제로 보인다. 일단 이렇게 장기와 단기로 나누면, 그 다음은 이 둘을 단계적으로 모두 시행해 볼 수 있을 것 같기도 하다. 즉, 단기적으로 순환조사를 시행하고, 장기적으로는 행정자료 기반의 센서스로 나아가는 방안이다. 그렇지만 이 방안은 이중 비용 부담에다, 중국에는 순환조사에서 개발한 방법을 폐기하게 되는 문제가 있다.

이 시점에서 순환조사를 표본조사에만 국한시켜 표본조사는 순환조사로 지속하되, 행정자료 기반의 센서스를 향해 전반적인 통계체제를 바꾸는 방안을 생각해 보자. 지금까지 표본조사는 전수조사와 함께 시행되면서 마치 둘은 뗄 수 없는 관계에 있는 것처럼 되어 왔다. 이 둘의 관계를 보면 실제 조사

내역은 크게 관계가 없지만, 차후에 다른 표본을 설계하는 데 이들은 함께 사용된다. 예를 들면 경제활동인구조사의 표본설계에 표본조사 자료가 사용되며, 이때 전수조사 자료는 모집단의 역할을 하고, 가중치를 부여하는 데 중요한 역할을 한다. 물론 완전한 행정자료 기반 센서스에서는 표본조사가 불필요하다. 단지 그 완전함을 확보하는 데 많은 시간이 들며, 그 사이에도 통계는 생산되어야 하기 때문에 여전히 표본조사가 필요하다. 이런 측면에서 당분간은 표본조사가 있어야 할 것으로 보인다.

적어도 당분간 표본조사가 있어야 한다는 것을 기정사실로 받아들이면, 이를 순환조사로 바꿀 것인가를 고려해야 한다. 앞서 이야기 방법론의 문제나 비용 등의 문제를 제외하더라도 또 다른 문제를 생각해 볼 수 있다. 표본조사가 60개월짜리 순환조사가 되면 현재 진행되는 다른 조사들과 어떻게 조화를 이룰지의 문제다. 경제활동인구조사와 비슷한 내역을 조사하는 것은 특히 문제가 될 수 있다. 이런 측면에서 다른 조사들에 대한 조정 작업이 있어야 할 것이다. 더욱이, 이런 모든 문제들을 고려하면, 표본조사의 시의성과 유용성에 관한 문제는 다른 방식의 소지역 통계로 대체가 가능한지도 검토해 보아야 할 사항이다.

이제 우리는 한국의 통계생산 전반을 놓고 고려하게 되었다. 다시 원점으로 돌아가야 할 때다. 다시 한번 현재의 상황을 그려보면, 행정에서 주민등록자료의 정비와 활용은 행정자료 기반의 센서스에 대한 요구로 이어지고 있다. 이를 거부할 명분이나 대안은 없다. 따라서 우리는 행정자료의 이용을 확대하며 행정자료 기반 센서스로 나아가야 한다. 행정에서 주민등록이 차지하는 역할의 비중과 활용 가능성이 소지역 통계에 대한 필요성보다는 훨씬 더 범위가 넓고 파급효과도 크기 때문이다.

이를 전제로 이제 통계청은 행정자료를 이용하기 위해 개별 행정부처에 자료를 요구하는 수준에서 벗어나 더 공격적으로 행정자료를 통계적 목적으로 활용할 수 있도록 행정업무의 설계와 집행에 한 부분을 담당할 필요가 있다. 어차피 개별 행정부처에서도 행정을 통한 통계집계가 점점 더 필요해지기 때문이다. 이를 발굴하고 알려주는 것도 통계청의 역할이다. 본 연구진은 통계청이 앞으로 이러한 역할을 하지 않으면 다른 부처가 이를 대신할

것이며 결국 그들이 필요한 통계기능을 가져가고 통계청의 역할은 줄어들 것이다. 예를 들면, 행정자치부와 정보통신부는 행정정보공유센터를 만들어 자신들의 정보를 연계하는 작업을 시작했다. 또 건설교통부와 행정자치부는 부동산정보관리센터를 통해 종합부동산세를 책정하기 위해 각종 대장을 연계하기 시작했다.

이런 측면에서 보면, 행정자료 기반의 센서스는 선택이 아니라 반드시 가야할 수밖에 없는 길이다. 표본조사를 순환조사로 할 것인가 아닌가는 그에 비하면 훨씬 작은 규모의 논의다. 따라서 수요가 있고 충분히 비용을 지불할 수단 있다면 시행할 수도 있는 것으로, 의사결정자들의 판단에 달린 문제다. 이에 비해 행정자료 기반의 센서스는 의사결정의 문제가 아니다. 이것은 시대의 주어진 상황이며, 통계청은 그에 비추어 미래에 설 자신들의 역할을 정립하는 비전을 갖는 것이 중요하다.

통계의 필요성에 비추어, 행정자료 기반의 자료 연계는 통계청이 주도하는 것이 가장 바람직하다. 그렇지 않으면 각종 행정정보가 개인수준에서 연결되어 개인을 통제하는 수단이 될 수 있기 때문이다. 이럴 경우 국민들이 행정정보의 연계 자체를 거부할 수도 있어 여러 행정자료들을 연계하는 시기가 매우 늦춰질 수 있다. 통계청은 개인수준에서 연계된 개인정보 자체에는 관심이 없으며, 이들을 집계하여 국가와 사회에 필요한 통계를 생산하는데 관심이 있다. 따라서 개인정보를 관리하는 하위 이를 개인수준의 정보로 활용하지 말아야 하는 서로 어긋날 것 같은 역할에는 통계청이 가장 적합하다. 각종 개인정보들을 개인수준에서 연계하여 종합적인 개인정보의 집합체를 구축하지만 이를 반드시 집계하여 통계로만 외부에 제공하는 것이 통계청 본연의 임무이기 때문이다. 이런 차원에서, 통계청은 개인정보의 연계와 이의 보호를 주도함으로써 개인정보보호에 민감하면서 동시에 그러한 자료를 모아 행정에 필요한 통계를 생산하고, 국민에게 국가통계를 제공하는 중심 역할을 선도해 가야 할 것이다.

제9장. 참고문헌

국문

- 강영옥. 2006. 『서울시 GIS 현황 분석평가 및 3단계 GIS기본계획 수립연구』. 서울시정개발연구원.
- 강영옥·이주일. 2005. 『건축물보 정비방안』. 서울시정개발연구원.
- 강영옥·윤은주·원종석·정재희. 2007. 『GIS기반 소지역 통계집계 공표구역의 획정 및 관리방안 연구』. 건설교통부.
- 건설교통부. 2004. 『개별주택가격 조사·산정 업무요령』.
- 건설교통부. 2005. 『건축물대장 정비 및 검수 지침』.
- 건설교통부·솔리테오. 2005. 『건축물대장 정비 중장기 사업계획(안)』.
- 국가예산정책처. 2006. 『인구주택총조사사업평가』.
- 김근영·김순관·이신혜. 2003. "Comparing Census Data with Resistration Data with Respect to Population and Household in the Seoul Metropolitan Region." 『국토계획』 38(1): 237-249.
- 김설희·강계화. 2005. "GIS DB에 연계시킨 행정자료를 활용한 현장조사 품질제고 방안." 조사연구학회 춘계학술대회 발표문.
- 김형석. 1999. "인구주택 총조사의 문제점과 개선방안 : 행정 등록통계와의 접목 방안에 대하여." 미발간 통계청 자료.
- 김형석·양경진·강영민. 2003. "인구센서스와 행정자료의 활용." 미발간 통계청 자료.
- 내무부·한국전산원·벽산엔지니어링. 1998. 『건축물대장 관리시스템 개발보고서』.
- 변필성·김광익·안재학. 2006. "소지역통계 공표구역 사례: 잉글랜드·웨일즈의 Output Area." 『국토』 10월호. 국토연구원.
- 서울시정개발연구원. 2001. 『서울시 도시계획 정보관리시스템 기초데이터 구축연구』.
- 서울시정개발연구원. 2002. 『모바일 GIS 적용에 관한 연구』.
- 서울시정개발연구원. 2003. 『서울시 모바일 행정시스템 도입을 위한 기본연구』.
- 서울특별시, 2001, 『서울시 도시계획 정보관리시스템 기초데이터

- 구축연구』 서울시정개발연구원.
- 서울특별시. 2005a. 『재산세 교육교재』. 세무과.
- 서울특별시. 2005b. 『위법(무허가) 건축물 지도점검계획』. 주택국 내부자료.
- 서울특별시. 2005c. 『개별주택가격 공시업무 편람』.
- 솔리데오시스템즈. 2004. 『용인시 건축물대장 전산자료정비 전산화 구축사업』.
- 안경애. 2004. “전국 건축물대장 DB 구축.” 「디지털 타임스」. 8월 30일.
- 양경진. 2003. “행정자료 활용 주택DB 구축 방안.” 「2005년 인구주택총조사 준비 제2차 워크숍」.
- 윤혁경(편저). 2003. 『건축법 조례 해설』. 기문당.
- 이건. 2007. “읍면동 수준에서 인구주택총조사와 주민등록의 집계인구 비교.” 『통계』. 33(1): 37-54.
- 이건·변미리·이명진·서우석. 2004. 『주거부문 행정자료의 인구주택 총조사 현장적용방안』. 통계청.
- 이건·이명진·서우석·변미리. 2006. “2005년 인구주택총조사의 조사구 설정 방법.” 『조사연구』 7(1):109-129.
- 이내성. 2006. “주민등록 행정자료를 활용한 인구통계 조사방법 연구.” 미발간 통계연구회 발표자료.
- 이명진·서우석·이건. 2006. “주거부문 행정자료의 인구주택 총조사 활용가능성에 대한 경험적 검토.” 『통계연구』 11(1): 57-87.
- 이삼식·박종서·하미영. 2002. 『환경 변화에 따른 인구주택총조사 개선방안』. 통계청·한국보건사회연구원.
- 이지연. 2007a. “2010년 인구주택 총조사 방법론 연구 기초자료.” 미발간 통계청 자료.
- 이지연. 2007b. “2005년 인구주택총조사의 범위오차평가와 지역별 차이 연구.” 한국인구학회 춘계학술대회 발표문.
- 임명선. 2003. “2005년 총조사 조사구설정방법 개선방안.” 「2005년 인구주택총조사 준비 제2차 워크숍」.
- 전라북도. 2006. 『2005 인구주택총조사결과 활용성제고방안 연구요약서』. 통계청. 2002. 『2000 인구주택총조사 CD』.
- 통계청. 2003a. 『기초단위구 현지확인지침서』.
- 통계청. 2003b. 『인구주택총조사 기초단위구를 이용한 통계적 지역구분에 관한 연구』.

- 통계청. 2004a. 『2005 인구주택총조사 조사지침서』 .
- 통계청. 2004b. 『2005 인구주택총조사 기본계획안(요약)』 .
- 통계청. 2004c. 『지역통계 생산을 위한 도시화지역 설정 연구』 .
- 통계청. 2005. 『2005 인구주택 총조사 조사지침서』 .
- 통계청. 2006a. 『2005 인구주택총조사 종합평가보고서』 .
- 통계청. 2006b. 『도시화지역 획정을 위한 시스템 개발』 .
- 통계청. 2006c. 『통계지리정보시스템 구축을 위한 ISP/BPR 사업』 .
- 통계청. 2007. 『2010 인구주택총조사 추진기본계획 (잠정안)』 .
- 통계청 인구조사과. 2003. 「2005년 인구주택총조사를 위한 자료이용기관 의견수집조사결과」 . 내부자료.
- 한국건설기술연구원. 2002. 『모빌GIS 기술을 이용한 지자체 지리정보 활용방안 연구보고서』 .
- 한국인구학회. 1998. 『행정수요 기준설정을 위한 새로운 인구개념의 정립 연구』 .
- 한국전산원. 1996. 『건축물대장의 전산화를 위한 서식표준화(안) 개발보고서』 .
- 행정자치부. 2004. 『읍면동 기능전환 관련 시군구 이관사무 처리 편람』 .
- 홍두승 · 이진 · 이명진 · 서우석 · 장원호. 2002. 『인구주택 총조사 개선방안에 관한 연구-행정자료의 인구주택 총조사 활용방안』 . 통계청.

영문

- Al, Pieter G. and Jan Willem Altena. 2000. Data security, privacy and the SSB. *Netherlands Official Statistics* Vol. 15 (Summer).
- Al, Pieter G. and Bart Bakker. 2000. "Re-Engineering Social Statistics by Micro-Integration of Different Sources." *N Netherlands Official Statistics* Vol. 15 (Summer).
- Alexander, Charles. H. 2000. "The American Community Survey and the 2010 U.S. Census." INSEE-Eurostat Seminar on Censuses after 2001.
- Alexander, Charles. H. 2002. "Still Rolling: Leslie Kish's 'Rolling Samples' and the Americann Community Survey." *Survey Methodology* 28:

35-41.

- Arts, Koo. 2000. "Matching Administrative Registers and Household Surveys." *Netherlands Official Statistics* Vol. 15 (Summer).
- Australian Bureau of Statistics. 2007. "The 2006 Australian Standard Geographical Classification and Census Geographic Areas - Diagram of Structures." <http://www.abs.gov.au/AUSSTATS/abs@.nsf/Latestproducts/2914.0Main%20Features225002006?opendocument&tabname=Summary&prodno=2914.0&issue=2006&num=&view>. 2007년 10월에 접속.
- Bierau, Dieter. 2000. "Testing a Register-Based Census." Insee-Eurostat Seminar on the Censuses after 2001. Paris.
- Borchsenius, Lars. 1996. "From a Conventional to a Register-Based Census of Population." Working Paper No.9. SCECE Work Session on Registers and Administrative Records in Social and Demographic Statistics. Geneva.
- Borchsenius, Lars. 2000. "From a Conventional to a Register-based Census of Population." INSEE-Eurostat Seminar on the Censuses after 2001. Paris.
- Borchsenius, Lars. 2006. "New Development in the Danish System for Access to Microdata." United Nations Statistical Commission and Economic Commission for Europe Conference of European Statisticians. Geneva.
- Bruhn, Åke. 2001. "The 2005 Population and Housing Census in Sweden will be Totally Register-Based." Symposium on Global Review of 2000 Round of Population and Housing Censuses. Statistics Division, Department of Economic and Social Affairs, UN Secretariat, New York, 7-10 August 2001.
- Canberra Group, The. 2001. *Expert Group on Household Income Statistics: Final Report and Recommendation*. Ottawa. <http://www.lisproject.org/links/canberra/finalreport.pdf> 2007년 11월 10일 접속.
- CES. 2003. "Writing Beneficial Statements for Projects Accessing Confidential Data." <http://www.ces.census.gov/index.php/ces/>

[researchguidelines](#). 2007년 11월 10일 접속.

- Choe, Minja Kim. 2000. " Electronic Microdata of the Censuses of the Republic of Korea at the East-West Center, University of Hawaii." Pp. 123-140 in *Handbook of International Historical Microdata for Population Research* (Ed.) Patricia Kelly Hall, Robert McCaa, and Gunnar Thorvaldsen. Minnesota, MN: Minnesota Population Center.
- Durr, Jean-Michel and J. Dumais. 2002. "Redesign of the French census of Population." *Survey Methodology* 28: 43-49.
- Edmonston, Barry and Charles Schultzd (Ed.). 1995. *Modernizing the U.S. Census*. Washington D.C.: National Academy Press.
- Eppmann, Helmut, Sonja Krügener, and Josef Schäfer. 2006. "First German Register Based Census in 2011" *Allgemeines Statistisches Archiv* 90(3): 465-482.
- European Commission. 2002. "Commission Regulation (EC) No 831/2002 of 17 May 2002: Implementing Council Regulation (EC) No 322/97 on Community Statistics, concerning Access to Confidential Data for Scientific Purposes." *Official Journal of the European Communities* 18.5.2002: 7-9.
- Gaasemyr, Svein (Statistics Netherland). 1999. "Contents and use of Register-based Job Files." Working paper No. 32. Joint ECE/Eurostat Work Session on Registers and Administrative Records in Social and Demographic Statistics. Geneva.
- Hall, Kelly, Robert McCaa, and Gunnar Thorvaldsen. 2000. *Handbook of International Historical Microdata for Population Research* (Ed.) Patricia Minnesota, MN: Minnesota Population Center.
- Harala, Riitta and Jari Nieminen. 1998. "Use of the Building and Dwelling Register in the Production of Statistics." *Statistics Finland*.
- Harala, Riitta and Reinikainen Anna-Leena. 1996. "Confidentiality in the Use of Administrative Data Sources." *Statistical Journal of the United Nations ECE* (1996): 361-368.
- Hawkers, William J. 1986. "Census Data Quality - A User's View." *Journal of Official Statistics* 2(4): 531-544.
- Hendricks, Coen, Paul Inge, and Harald Utne. 2000. "Establishing a

- Dwelling Register in Norway: Completing the Register System." Insee-Eurostat Seminar on Census after 2001. Paris.
- INSEE. 2006. "The Population Censuses in France : from General Censuses to 'Rolling Census'." UN Economic and Social Council.
- INSEE. 2004. "The New French Rolling Census." UNECE Seminar on New Methods for Population Censuses.
- Jensen, Lothar. 2000. "Opening Speech." Insee-Eurostat Seminar on the Censuses after 2001. Paris.
- Juristisches Internetprojekt Saarbrücken. 2001. Gesetz zur Vorbereitung eines Registergestützten Zensus.
- Kish, Leslie. 1990. "Rolling Samples and Censuses (with discussion)." *Survey Methodology* 16: 63-79.
- Kish, Leslie. 1998. "Space/Time Variations and Rolling Samples." *Journal of Official Statistics* 14: 31-46.
- Kish, Leslie. 1999. "Cumulating/Combining Population Surveys." *Survey Methodology* 25: 129-138.
- Laihonen, Aarno. 1999. "Development of the Use of Administrative Data in Population and Housing Censuses in Europe." Working Paper No. 6. Joint ECE/Eurostat Work Session on Registers and Administrative Records in Social and Demographic Statistics. Geneva.
- Laihonen, Aarno. 2000. "2001 Round Population Censuses in Europe." Insee-Eurostat Seminar on Census after 2001. Paris.
- Martin, David. 1998b. "Optimizing Census Geography: the Separation of Collection and Output Geographies." *Geographical Information Science* 12(7): 673-685.
- Martin, David. 2002. "Geography for the 2001 Census in England and Wales." *Population Trends* 108: 7-15.
- Martin, David. 2004. "Neighbourhoods and Area Statistics in the Post 2001 Census Era." *Area* 36.2: 136-145.
- Myrskylä, Pekka. 1999. "New Statistics Made Possible by the Use of Registers." Working Paper No. 8. Joint ECE/Eurostat Work Session on Registers and Administrative Records in Social and Demographic Statistics. Geneva.

- Myrskylä, Pekka. 2000. Effects of Changes to Registers: Experiences from Finnish Register-based Censuses." Insee-Eurostat Seminar on Census after 2001, Paris.
- Nacional de Estadística (Spain). 1999a. "Utilization of Administrative Registers in the 2001 Spanish Demographic Censuses." Working paper No. 27. Joint ECE/Eurostat Work Session on registers and Administrative Records in Social and Demographic Statistics. Geneva.
- Nacional de Estadística. 1999b. "The 'Padron Continuo'." Working paper No. 28. Joint ECE/Eurostat Work Session on Registers and Administrative Records in Social and Demographic Statistics. Geneva.
- Nordberg, Lief, Irmeli Penttilä, and Susanna Sandström. 2001. "A Study on the Effects of Using Interview versus Register Data in Income Distribution Analysis with an Application to the Finnish ECHP-survey in 1996." CHINTEX Working Paper #1.
- Nordholt, E, Hartgers, M and Gircour R. 2004. The Dutch Virtual Census of 2001: Analysis and Methodology, Netherlands Central Bureau of Statistics Monograph. Statistics Netherlands: Voorburg/Heerlen.
- Office for National Statistics(ONS). 1998. *2001 One Number Census*. London.
- Office for National Statistics(ONS). 2001. *Census 2001: A Guide to the One Number Census*. <http://www.statistics.gov.uk/nsbase/census2001/pdfs/oncguide.pdf>. 2007년 8월에 접속.
- Office for National Statistics(ONS). 2003. "Alternative to a Census: Review of Interantional Approaches." *Census Strategic Development Review*. <http://www.statistics.gov.uk>. 2007. 10월 20일 접속.
- OPCS. 1992. "Policy Evaluation and Review Project Annex 5: Rolling Censuses." Unpublished International Research Paper.
- Puderer, Henry. 2001. "Introducing the Dissemination Area for the 2001 Census: an Update." *Geography Working Paper Series* No. 2000-4. Statistics Canada.
- Schaafsma-Harteveld, Berna (Statistics Netherland). 1999. "Disablement

- Benefits: Combining Survey Data with Register Records." Working paper No. 20. Joint ECE/Eurostat Work Session on registers and Administrative Records in Social and Demographic Statistics. Geneva.
- Slagter, Herman (Statistics Netherland). 1999. "Compiling Structure of Earning Statistics using Existing Survey Data and Register Data." Working paper No. 7. Joint ECE/Eurostat Work Session on registers and Administrative Records in Social and Demographic Statistics. Geneva.
- Spieker, Finn. 1999. "Formation of Central Variables in a Decentralised Statistical System." Working Paper No. 25. Joint ECE/Eurostat Work Session on Registers and Administrative Records in Social and Demographic Statistics. Geneva.
- Statistisches Ämter. 2007. The Census Process 2011. http://www.statistik-portal.de/Statistik-Portal/en/Zensus/en_zeitAblauf.asp. 2007년 10월에 접속.
- Statistisches Bundesamt Deutschland. 1992. "Considerations of Alternatives to Censuses and Census-type Statistics: the Case of Germany." Working Paper No. 13. SCECE Work Session on Population and Housing Censuses. Geneva.
- Statistisches Bundesamt Deutschland. 2002. "Der Test eines Registergestützten Zensus." <http://www.destatis.de>. 2004년 9월에 접속.
- Statistisches Bundesamt Deutschland. 2007. *Microdata*. <http://www.destatis.de>. 2007년 10월에 접속.
- Statistics Canada. 2004. *2001 Census Dictionary*. Ottawa: Ministry of Industry.
- Statistics Finland. 2006. "Access to Enterprise Data at Statistics Finland: Introducing the Practices of the Research Laboratory." OECD Conference: Assessing the Feasibility of Microdata Access, Luxembourg 26-27 October 2006.
- Statistics Netherland. 2006. "The Dutch Virtual Census of 2001: A Register-Based Approach Combined with Survey Information." Paper

- Presented at the Conference of European Statisticians for UNESCO.
Paris.
- Statistics Sweden. 1999. "Registers in Official Statistics: A Swedish Perspective." Working Paper No. 14. Joint ECE/Eurostat Work Session on Registers and Administrative Records in Social and Demographic Statistics. Geneva.
- Task Force on Statistics for Rural Development and Agriculture Household Income. 2005. "The Agricultural Household - Concepts and Definitions (Part 1)." Joint UNECE/EUROSTAT/FAO/OECD Meeting on Food and Agricultural Statistics in Europe. Rome.
- Trewin, Dennis. 2006. *Statistical Geography Vol.1: Australian Standard Geographical Classification*. Australian Bureau of Statistics.
- U.S. Census Bureau. 2004. "Meeting 21st Century Demographic Data Needs - Implementing the American Community Survey." Report 8: Comparison of the American Community Survey Three-Year Averages and the Census Sample for a Sample of Counties and Tracts. Washington D. C.
- U.S. Department of Commerce • Economics and Statistics Administration • Bureau of the Census. 2000. "United States Census 2000 Participant Statistical Areas Program Guidelines."
- Utne, Harald. 2005. *The Population and Housing Census Handbook 2001*. Statistics Norway. <http://www.ssd.no>.
- van der Laan, Paul. 2000. "The 2001 Census in the Netherlands Integration of Registers and Surveys." Insee-Eurostat Seminar on the Censuses after 2001. Paris.
- Vliegen, Mathie and Huib Van de Stadt (Netherlands Central Bureau of Statistics). 1989. "Is a Census Still Necessary? Experiences in the Netherlands." Working Paper No. 5. SCECE Seminar on the Relevance and Importance of Population and Housing Census Data. Wiesbaden, Germany.
- Zeithaml, Valerie A. and Mary Jo Bitner. 1996. *Services Marketing*. New York: Mcgraw Hill.

일문

西村善博. 2006. “센서스와 統計調査의 變容.” 『사회과학으로서의 통계학』.
제4집: 63-73. 産業統計研究社.

<부록 1> 온라인 설문지



인구 센서스 이용자 조사

안녕하십니까?

조사에 참여해 주셔서 대단히 감사합니다.

통계청에서는 보다 질 좋은 통계자료를 제공하기 위하여 한국조사연구학회(연구책임자 : 이건 서울시립대 도시사회학과 교수)와 [인구주택총조사 자료이용실태 및 만족도 조사]를 실시하고 있습니다.

귀하께서 응답해 주신 내용은 통계청이 5년마다 실시하고 있는 인구주택총조사의 이용효율성과 향후 개선방향을 파악하는 귀중한 자료가 될 것입니다.

본 조사의 내용은 맞고 틀리는 답이 있는 것이 아니며, 귀하께서 평소 생각하신 점을 솔직하게 말씀해 주시면 됩니다.

응답해 주신 모든 분들의 개인정보는 통계법 제 13 조에 의해 비밀이 보호되며 응답에 소요되는 시간은 약 10분 정도입니다.

응답해 주시는 분들 중 20명을 선정하여, 5만원 상당의 문화상품권(혹은 도서상품권 등)을 조사 종료 후에 보내드리도록 하겠습니다.

바쁘시더라도 잠시만 시간을 내어 협조해 주시면 감사하겠습니다.

☞ 본 조사에 관한 의문사항이나 궁금한 점이 있으시면 아래의 연락처로 문의해 주시기 바랍니다.

서울특별시 동대문구 전농동 시립대길 13
서울시립대학교 도시사회학과 대학원 남궁상/조민경/이정은
Tel. 02) 2210-XXXX
Email

▶ 다음



인구 센서스 이용자 조사

설문진행률 :  7%

Q1. 선생님께서는 다음 중 어디에 속하십니까?

- 중앙부처 공무원
- 지방자치단체 공무원
- 중앙부처 산하 연구원 소속의 연구원
- 지방자치단체 산하 연구원 소속의 연구원
- 기타()

▶ 다음



인구 센서스 이용자 조사

설문진행률 : 14%

Q2_2. 선생님께서는 다음 중 어느 기관에 소속되어 있으십니까?(시/군/구청 공무원이나 지자체 산하 연구원에 소속되어 있는 경우, 상위기관을 선택하여 주십시오.)

- | | | | |
|-----------------------------|-----------------------------|-----------------------------|-------------------------------|
| <input type="radio"/> 서울특별시 | <input type="radio"/> 부산광역시 | <input type="radio"/> 대구광역시 | <input type="radio"/> 인천광역시 |
| <input type="radio"/> 광주광역시 | <input type="radio"/> 대전광역시 | <input type="radio"/> 울산광역시 | <input type="radio"/> 경기도 |
| <input type="radio"/> 강원도 | <input type="radio"/> 충청북도 | <input type="radio"/> 충청남도 | <input type="radio"/> 전라북도 |
| <input type="radio"/> 전라남도 | <input type="radio"/> 경상북도 | <input type="radio"/> 경상남도 | <input type="radio"/> 제주특별자치도 |

Q3. 선생님께서 일하고 있는 부서나 팀명을 적어주십시오.

◉ 다음



인구 센서스 이용자 조사

설문진행률 : 21%

Q4. 선생님께서는 인구주택총조사 결과를 업무에 이용하신 적이 있습니까? 있다면, 어떤 통계 자료를 이용해 보셨습니까? 이용해 본 경험이 있는 자료들을 모두 선택하여 주십시오.

- 인구 - 성별, 나이, 혼인상태, 교육정도, 종교, 가구주와의 관계, 인구이동 등
- 교통 - 통근통학여부, 이용교통수단, 소요시간 등
- 경제활동 - 경제활동상태, 종사상 지위, 직업 등
- 가구 - 가구구성형태, 거주기간, 자동차 보유대수 등
- 주택 - 거처의 종류(단독주택/아파트 등), 주택소유여부, 소유형태 등
- 기타 - 이산가족, 특체적?정신적 활동제한, 고령자 생활비 현천
- 자녀 출산시기
- 현 시/군/구 거주사유, 이전시군구 이동사유
- 이용해 본적 없음

◉ 다음



인구 센서스 이용자 조사

설문진행률 : 28%

Q5. 선생님께서는 어떠한 형태의 인구주택총조사 결과를 이용하십니까? 이용해 보셨던 자료의 형태를 모두 선택하여 주시기 바랍니다.

- 통계청 인구주택총조사 보고서
- 통계청 홈페이지
- 통계청 보도자료
- CD롬 형태의 결과표보고서/원자료(Raw Data)
- 소속 지자체 홈페이지
- 기타()

Q6. 선생님께서는 인구주택총조사 통계 자료를 주로 어떤 용도로 활용하십니까?

- 정책수립 및 평가
- 연구·학술·학습활동
- 사업 및 경영계획 수립
- 표본조사의 모집단
- 교육자료 인용
- 기타()

다음



인구 센서스 이용자 조사

설문진행률 : 35%

Q7. 선생님께서는 인구주택 총조사 통계자료를 얼마나 자주 이용하시는 편입니까?

- 주 1회 이상
- 월 1회 이상
- 분기 1회 이상
- 반기 1회 이상
- 연 1회 정도
- 연 1회 미만

Q8. 인구주택총조사 통계에서 어떤 분야의 자료를 자주 이용하시는 편입니까? 아래에 제시되어 있는 분류표를 참고하셔서, 선생님께서 자주 이용하는 자료를 순서대로 3가지만 선택하여 주시기 바랍니다. 선생님께서 이용하시는 부분이 제시되어 있지 않으면 기타에 구체적으로 적어주십시오.

성별	나이
혼인상태	교육정도
종교	가구주와의 관계
인구이동	
기타 ()	

다음



인구 센서스 이용자 조사

설문진행률 : 42%

Q9. 인구주택총조사 통계자료에 대해 얼마나 만족하십니까? 선생님께서 이용해 보셨던 경험을 바탕으로 각 항목별로 응답해 주십시오.

속 성	전혀 그렇지 않다	대체로 그렇지 않다	보통 이다	대체로 그렇다	매우 그렇다
	1	2	3	4	5
1. 인구주택총조사 통계는 이용시 유의사항/관련개념/용어의 정의 등이 잘 되어 있다.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. 인구주택총조사 통계는 제공방법/편제가 필요한 통계정보를 얻는데 편리하게 구성되어 있다.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. 인구주택총조사 통계는 알고자 하는 충분한 정보를 제공하고 있다.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. 인구주택총조사 통계는 내가 하고 있는 업무에 도움이 많이 된다.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. 인구주택총조사 통계를 이용하는데 절차가 복잡하지 않고 편리하다.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. 인구주택총조사 통계는 적절한 시기에 공표 및 제공이 되고 있다.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

▶ 다음



인구 센서스 이용자 조사

설문진행률 : 50%

Q10. 인구주택총조사 통계자료에 대해 전반적으로 얼마나 만족하십니까? 선생님께서 앞서 항목별 평가내용을 고려하셔서 응답해 주십시오.

매우 불만족

불만족하는 편

보통

만족하는 편

매우 만족

▶ 다음



인구 센서스 이용자 조사

설문진행률 : 57%

Q11. 선생님께서 인구주택총조사 통계 자료를 이용하면서 느끼는 문제점은 무엇이라고 생각하십니까? 가장 불편했거나 문제라고 생각하셨던 순서대로 3가지만 선택하여 주십시오.

- 제공시기가 적절하지 않다
- 제공정보가 충분하지 않다
- 원자료 제공 절차가 불편하다
- 추가적인 정보를 얻는 방법에 대한 정보가 부족하다
- 인터넷 홈페이지에 있는 자료를 찾아보기가 불편하다
- 보고서 편제(편집과 형식)가 이용자들이 찾아보기에 불편하다
- 용어의 개념 설명, 이용 시 유의사항 등이 충실하게 수록되어 있지 않다
- 기타 ()

다음



인구 센서스 이용자 조사

설문진행률 : 64%

Q11_1. 구체적으로 어떻게 불편하셨습니까? 선생님께서 불편하다고 하셨던 각각의 항목에 대해 말씀해 주십시오.

- 1순위 : 제공정보가 충분하지 않다
- 2순위 : 인터넷 홈페이지에 있는 자료를 찾아보기가 불편하다
- 3순위 : 추가적인 정보를 얻는 방법에 대한 정보가 부족하다

다음



인구 센서스 이용자 조사

설문진행률 :  71%

Q12. 선생님께서는 인구주택총조사 통계자료에 대해 전반적으로 얼마나 만족하십니까? 선생님께서 앞서 항목별 평가내용을 고려하셔서 응답해 주십시오.

- 전혀 이용할 의향 없다
- 이용할 의향없는 편이다
- 보통이다
- 이용할 의향있는 편이다
- 이용할 의향 매우 많다

Q13. 선생님께서 인구주택총조사에 추가되었으면, 또는 좀 더 구체적으로 조사했으면 좋겠다고 생각하신 항목들이 있었습니까? 어떤 것들이 있습니까?
생각하셨던 것들을 모두 말씀해 주시기 바랍니다.

 다음

<부록 2> 유럽 국가들의 센서스 방식



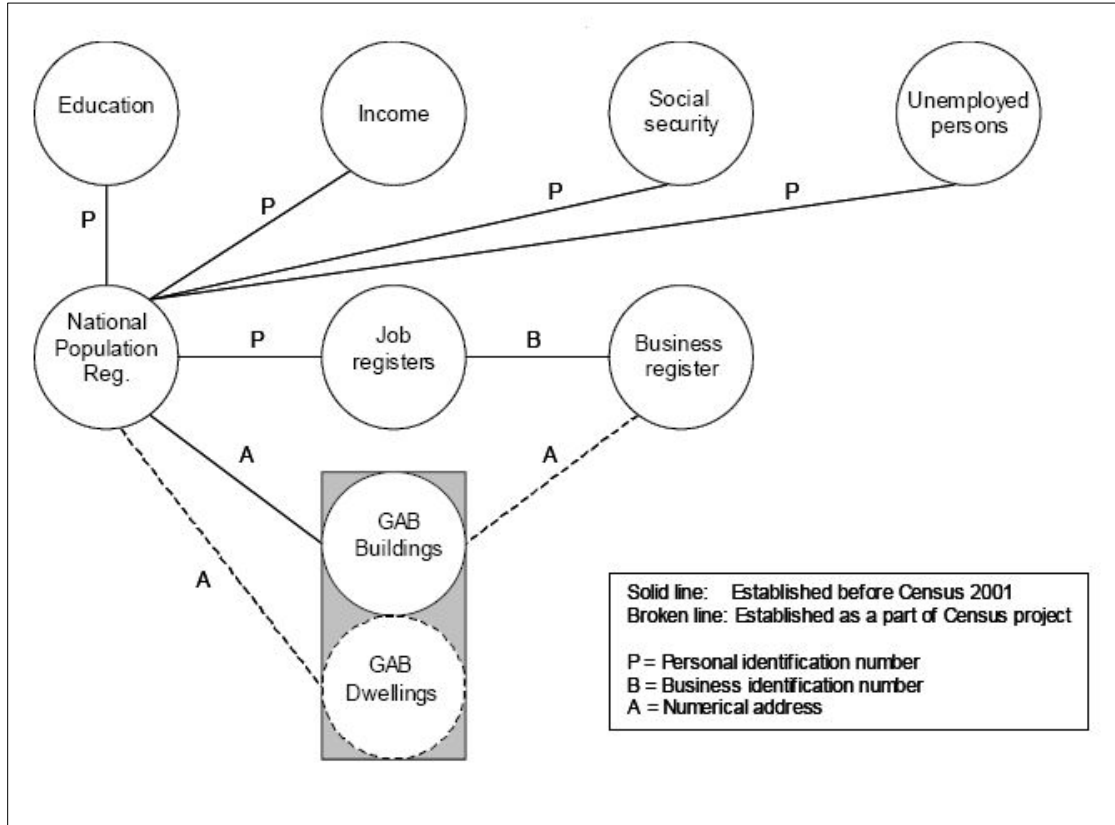
This is generally based on the latest census round performed. Where more recent information was available, it has been taken into account. Status: May 2006

*) Excluding Azores, Madeira, the Canary Islands and the French Overseas Territories and mini states.

출처 : http://www.statistik-portal.de/Statistik-Portal/en/Zensus/en_zensenA.asp

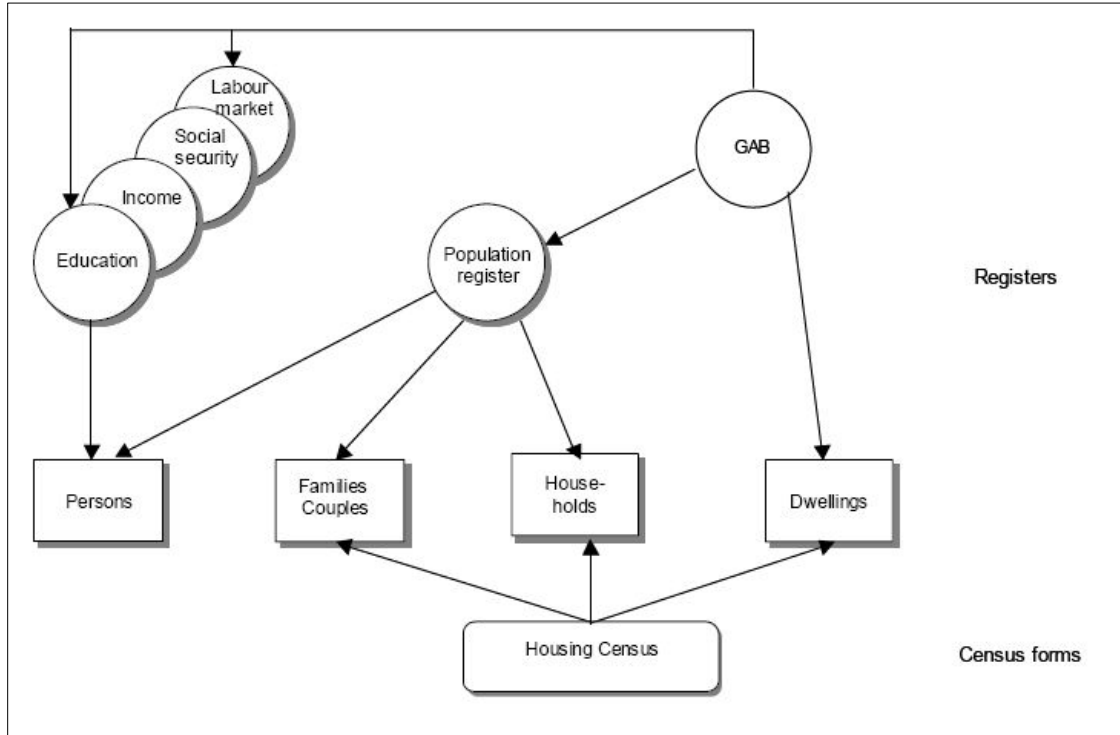
(독일 통계청 홈페이지) 2007년 10월 접속.

<부록 3> 노르웨이 2001년 센서스 행정자료 활용



(그림 부록3-1) 2001년 센서스 이전과 이후의 등록체계

출처 : Utne(2005: 6)



(그림 부록3-2) 2001년 센서스의 자료 출처
출처 : Utne(2005: 8)

<부록 4> 세대의 법적 정의

재정부 부동산세법 시행령 설명

□ 1세대란 본인 및 배우자와 동일한 주소에서 생계를 같이 하는 가족 (직계존비속, 형제자매)으로 구성하는 하나의 집단을 말함

○ 취학, 요양, 근무 형편에 따른 일시퇴거자도 1세대로 보아 합산함

<사례1> 혼인한 아들의 경우 분가하여 별도의 세대가 가능하나 부모와 동일 주소에서 생계를 같이하는 경우 예는 합산과세 함

* 대법원83누44(1983.4.26) : 생계를 같이 한다는 것은 반드시 주민 등록상 세대를 같이 함을 요하지 않으나 일상생활에서 볼 때 유무상통하여 동일한 생활자금에서 생활하는 단위를 의미한다고 해석함이 상당함

<사례2> 30세 이상의 유학중인 자녀가 아파트 1채를 소유하고 있으나, 부모와 생계를 같이하는 경우 1세대로 보아 합산과세 함

□ 배우자가 없더라도 다음의 경우는 1세대로 인정함

○ 30세 이상인 경우

○ 세대원이 당해 주택을 소유하면서 독립된 생계를 유지 할 수 있는 경우로서 국민 기초생활 보장법에 의한 최저생계비 수준 이상의 소득이 있는 경우. 다만, 미성년자는 제외(가족의 사망, 결혼 기타 재정경제부령이 정하는 사유로서 부득이하게 1세대를 구성하는 경우에는 소득기준 충족시 허용)

○ 배우자가 사망하거나 이혼한 경우

□ 부부 는 단독세대를 구성하는 경우에도 동일한 세대 로 봄

○ 본인과 배우자 는 주민등록상 별도세대를 구성한다 하더라도 1세대로 보아 합산 하여 과세 함 (대법원97누19465,1998.5.29 : 배우자는 그 배

우자라는 사실만으로 거주자와 1세대를 구성)

이혼 한 경우에는 합산하지 아니하나, 이혼 후 사실상 혼인 관계를 유지하는 위장이혼 인 경우에는 합산함

미성년자는 원칙 적으로 단독세대를 인정하지 아니함

다만, 생계를 유지할 수 있는 소득이 있는 경우로서

○ 가족의 사망, 결혼 등의 사유로 부득이하게 1세대를 구성 하는 경우는 단독세대를 인정함

4. 혼인·노부모 봉양을 위해 세대를 합가한 경우 어떻게 과세되나

합가 한 경우 1세대이므로 원칙 적으로 종부세를 합산과세 하나

○ 합가한 날로부터 2년간 은 합가전의 상태를 각각 1세대 로 보아 종부세 과세여부를 판단함

다만, 2006. 6. 1 이전 에 이미 혼인 또는 노부모 봉양 을 위해 합가한 경우

<부록 5> 주민등록에 관한 법규

주민등록법시행령

- ③ 주민등록표 등 주민등록관계서류의 주소는 특별시·광역시·도, 시·군·자치구, 구(자치구가 아닌 구를 말한다), 읍·면·동, 리, 지번의 순으로 기록한다. 이 경우 동·리는 법정 동·리의 명칭으로, 「주택법」의 규정에 의한 공동주택의 경우에는 지번 다음에 건축물관리대장등에 의한 공동주택의 명칭과 동·호수를 기록한다.

[개정 2003.11.29, 2005.3.31, 2006.9.6] [[시행일 2006.9.25]]

[본조제목개정 2006.9.6] [[시행일 2006.9.25]]

제4조 (세대에 속하지 아니한 군인의 등록) 법 제6조제2항의 규정에 해당하는 군인으로서 어느 세대에도 속하지 아니한 군인에 대하여는 그 군인의 본적지의 시장·군수 또는 구청장이 등록하여야 한다.

주민등록법시행규칙

제6조 (거주지이동에 따른 관련자료의 정리 등)

- ① 전입신고를 받은 시장·군수 또는 구청장은 영 제20조 및 영 제21조의 규정에 의하여 이송받은 주민등록표를 지체없이 정리하여야 한다.
- ② 시장·군수 또는 구청장이 제1항의 규정에 의한 정리를 한 때에는 영 제6조제3항의 규정에 의한 세대명부·주민등록전출자명부 및 주민등록전입자명부(이하 "세대명부등"이라 한다)를 작성하여야 한다.

<부록 6> 지방자치단체에 관한 규정

지방자치법

제2조 (지방자치단체의 종류)

① 지방자치단체는 다음의 두 가지 종류로 구분한다.

1. 특별시, 광역시, 도, 특별자치도
2. 시, 군, 구

② 지방자치단체인 구(이하 "자치구"라 한다)는 특별시와 광역시의 관할 구역 안의 구만을 말하며, 자치구의 자치권의 범위는 법령으로 정하는 바에 따라 시·군과 다르게 할 수 있다.

③ 제1항의 지방자치단체 외에 특정한 목적을 수행하기 위하여 필요하면 따로 특별지방자치단체를 설치할 수 있다.

④ 특별지방자치단체의 설치·운영에 관하여 필요한 사항은 대통령령으로 정한다.

제3조 (지방자치단체의 법인격과 관할)

① 지방자치단체는 법인으로 한다.

② 특별시, 광역시, 도, 특별자치도(이하 "시·도"라 한다)는 정부의 직할(直轄)로 두고, 시는 도의 관할 구역 안에, 군은 광역시나 도의 관할 구역 안에 두며, 자치구는 특별시와 광역시의 관할 구역 안에 둔다.

③ 특별시 또는 광역시가 아닌 인구 50만 이상의 시에는 자치구가 아닌 구를 둘 수 있고, 군에는 읍·면을 두며, 시와 구(자치구를 포함한다)에는 동을, 읍·면에는 리를 둔다.

④ 제7조제2항에 따라 설치된 시에는 도시의 형태를 갖춘 지역에는 동을, 그 밖의 지역에는 읍·면을 두되, 자치구가 아닌 구를 둘 경우에는 그 구에 읍·면·동을 둘 수 있다.

<부록 7> 인구주택총조사에서 GIS 활용 방안

제1장. 조사구 설정과 GIS

제1절. 전통적 방식의 조사구 설정업무¹⁾

1. 개요

인구주택총조사는 연도의 끝자리가 0과 5인 해에 전국의 거주인구와 주택을 대상으로 11월 1일부터 15일 사이에 실시한다. 비록 보름 동안에 수행되지만 워낙 대규모 조사이면서도 단기간에 조사를 완료해야하기 때문에 많은 준비가 필요하며, 통상 인구주택총조사에 대한 준비는 2년여에 걸쳐 진행된다.

2005년도 인구주택총조사의 경우 크게 5 단계로 나뉘어 진행되었다. 그 단계는 사전 준비단계(2004년 1월~12월), 본조사 추진단계(2005년 1월~11월), 자료 처리·심사단계(2005년 11월~2006년 4월), 결과 분석 및 공표단계(2005년 12월~2006년 12월), 최종평가단계(2006년 6월~2007년 3월)이다(통계청, 2004b).²⁾

아래 (그림 부록7-1)은 조사와 관련된 준비과정을 종합하여 총조사의 일정에 맞추어 제시한 것이다. 인구주택총조사를 준비하는 과정 가운데 가장 중요한 작업 중 하나가 조사구를 설정하는 것이다. 인구주택총조사는 전국의 모든 가구를 15일 동안에 방문하여 조사하기 위해 약 14만 명의 조사원을 임시로 고용한다. 이렇듯 많은 조사원을 한꺼번에 동원하기 때문에 원활한 조사관리가 매우 중요하다. 이를 위해 통계청은 전국의 모든 지역을 일정한 규모의 가구 수를 기준으로 나누어 표준화하여 조사구를 정한다. 조사구는 통상 60 가구를 기준으로 나눈다. 이렇게 설정한 조사구를 지역별로 지도에 표시하여 조사구 지도(조사구요도)를 만든다. 조사구와 조사구 지도는 총조사 준비단계 전 과정에서 여러 차례의 자료 대조 및 실제 현장의 확인을 통

1) 이건, 이명진, 서우석, 변미리(2006)에서 발췌

2) 2010년 인구주택총조사 추진일정은 <부록 8> 참조

해 수정·보완된다.



(그림 부록7-1) 인구주택총조사 조사부분의 흐름
출처 : 통계청 내부자료 및 통계청(2004a)

본조사 직전 약 1 주일 동안에 준비조사가 있다. 준비조사는 말 그대로 총조사를 준비하는 조사이다. 이때 조사원이 조사구 지도를 갖고 조사구를 방문하여 조사구 지도에 표시된 도로 및 건물과 그 건물에 사람이 사는지 여부와 건물 안의 가구 수 등을 확인한다. 이를 토대로 통계청은 조사구와 조사구 지도를 수정하여 확정하고, 가구명부를 작성한다. 이렇게 만들어진 가구명부와 조사구 지도는 본조사에서 가장 중요한 조사도구로, 조사원은 이들을 바탕으로 가구를 방문하여 조사표를 작성한다.

2. 거처 파악과 조사구 설정

가. 거처

조사구 설정에서 가장 먼저 고려할 사항은 사람이 사는 건물을 파악하는 것이다. 사람들은 다양한 형태의 건물에서 살고 있다. 단독주택이나 아파트와 같은 공동주택의 경우는 거주여부가 명확한데 반해, 상가나 창고 등 거주 이외의 목적으로 지어진 건물에도 사람들이 거주하는 경우가 종종 있어 이

러한 건물들을 조사할 것인지에 대해서도 판단을 해야 한다. 이와 같이 복잡한 거주의 현황을 파악하기 위해 통계청에서는 사람이 살고 있는 장소를 거처라는 개념으로 파악한다. <표 부록7-1>은 통계청이 분류한 거처의 종류이다.

<표 부록7-1> 인구주택총조사의 거처 구분

주택	1) 단독주택(일반단독, 다가구단독, 영업겸용단독) 2) 아파트 3) 연립주택 ┆ 공동주택 4) 다세대주택 ┆ 5) 비거주용 건물 내 주택(상가, 공장, 점포, 여관 등에 주택 시설)
주택 이외의 거처	1) 오피스텔(사무실겸용 주택) 2) 호텔·여관 등 숙박업소의 객실 3) 기숙사 및 특수사회시설 4) 판잣집, 비닐하우스, 움막 5) 기타

출처 : 통계청(2004a)

1970년대까지는 우리나라 사람들의 대부분이 단독주택에 살았지만 최근들어 아파트가 크게 증가하여 모든 주택의 반 가량을 차지하게 되었다. 사람들이 다양한 형태의 주택에서 살고 있기 때문에 조사현장에 대해 상당히 잘 알고 있지 않으면 실제 조사에서 조사원이 거처를 제대로 파악하기란 그다지 쉽지 않다. 1995년도 인구주택총조사까지는 거주 현황을 잘 알고 있는 읍면동사무소에서 통계를 담당한 공무원이 거처를 파악할 수 있는 정보를 제공함으로써 조사구 설정이 수월하였다. 그러나 1999년부터 시행된 행정자치부의 읍면동 기능전환 사업으로 통계기능이 시군구로 이관되면서 기초자치단체에 통계기능이 없어 총조사에서 거처 파악이 어려워졌다. 다만 2000년 인구주택총조사 때는 아직 읍면동의 기능 이전이 시작 된지 얼마 지나지 않아 지방자치단체의 통계담당자를 어느 정도 활용할 수는 있었다. 그렇지만 2005년도 인구주택총조사부터는 읍면동 단위의 지방자치단체로부터 거처 파악 등 통계관련 업무에 관한 공식적 협조를 받을 수가 없는 실정이었다.

이런 행정체제의 변화에 따라 통계청은 거처 파악을 위한 다양한 방법을

모색하고 있다. 통계청은 2000년 인구주택총조사에서 파악된 건물과 거처에 대한 자료를 보유하고 있기 때문에 이 자료를 활용할 수 있다. 그렇지만 2000년 이후 건물이 없어지거나 새로운 건물이 들어선 경우는 거처의 차이가 발생한다. 건물이 없어졌을 때에는 실제 조사를 통해 거처를 지우게 된다. 신축 건물이 있을 때는 건축물대장과 건축인허가대장을 이용하여 거처를 확인하고 있다. 그러나 건축물 대장은 내용의 오류나 누락이 많아 사용하기 어려우며, 시군구마다 자료의 정확도가 달라 이에 전적으로 의존할 수 없다(안경애, 2004). 이런 경우에는 최근 작성된 수치지도를 이용하여 건물과 거처를 파악하기도 한다. 또한 아파트의 경우는 한국감정원의 감정평가 기초자료를 활용하여 거처를 파악하기도 한다(양경진, 2003).

나. 조사구

거처확인을 바탕으로 거처별로 몇 가구가 사는지를 파악하여 60가구 내외로 조사구를 (가)설정한다. 조사구에는 일반조사구와 특별조사구가 있으며, 세부적으로는 아파트조사구, 보통조사구 등이 있다(<표 부록7-2> 참조).

<표 부록7-2> 조사구 분류

대분류	소분류	비고
일반조사구	아파트조사구	60 가구 기준으로 지역을 분할해 설정
	보통조사구	
	섬조사구	
	기숙사시설조사구	개별시설단위로 설정
	특수사회시설조사구	호텔과 외국인 거주지역
관광호텔 및 외국인 거주지역조사구		
특별조사구	재외주재공관(외무부)	조사원으로는 조사가 불가능한 지역
	교도소, 소년원 등(법무부)	
	육·해·공군(국방부)	
	전투경찰대 등(경찰청)	
	전투경찰대 등(해양경찰청)	

출처 : 통계청(2002)

2005년 인구주택총조사에서는 읍면동에서 조사구를 설정할 수 없기 때문에 조사구 가설정 및 설정 작업을 모두 통계청에서 해야 했다. 그렇지만 가

설정된 조사구와 변동 조사구를 확인하여 조사구를 수정·보완하기 위해서는 지방자치단체의 협조를 받는 것은 불가피하다. 이에 조사구 설정에서는 시군구 통계담당 직원의 업무 협조를 받았다(임명선 2003). 이런 방식으로 2004년 하반기에 가구 수를 파악하였고, 2005년 상반기에 조사구를 설정하였다. 한 가지 참고할 사항은 기초단위구와 관련된 것이다. 한때 조사구 설정에 기초단위구를 활용할 계획이 있었으나 시험조사를 통해 살펴본 결과 조사구별 가구수 편차와 변동, 조사구 경계설정의 어려움 등으로 이번 총조사에서는 기초단위구를 활용하지 않았다(임명선 2003). 그러나 앞으로 안정된 기초단위구가 확보되면 이의 활용도가 높아질 것으로 보인다.

조사구 설정은 조사할 지역의 구획을 나누는 것으로 공간적으로 고려해야 할 사항들이 많다. 조사구는 조사구 안의 모든 거처를 빠지지 않고 가능하면 짧은 시간 안에 방문할 수 있도록 이동경로가 단순·명료하며, 장애물이 없어야 하고, 총 이동거리가 길지 않게 설정하는 것이 바람직하다. 따라서 조사구 설정에는 도로와 건물이 표시된 지도가 필수적인 도구이다. 2000년 인구주택총조사부터 통계청은 행정자치부의 수치지도를 활용하여 GIS(Geographical Information System)를 기반으로 조사구 설정용 기본도 및 조사구 지도를 작성하였다.

통계청은 2000년 인구주택총조사에서 수치지도를 사용하여 조사 과정에서 조사구 지도를 수정할 수 있었다. 이 지도는 그 당시로는 상당히 정확하며, 적어도 조사구 설정을 위해서는 가장 정확하다고 할 수 있다. 앞서 설명한 거처 확인 작업과 마찬가지로, 통계청은 이 자료에 2000년 11월 이후에 새로 제작된 수치지도를 붙이고, 더 나아가 허물거나 새로 건축한 건물을 파악하여 지도를 수정하였다. 이때 허물거나 새로 신축한 건물에 대해서는 행정자치부와 지방자치단체가 보유하고 있는 건축인허가대장과 건축물대장의 도면을 사용하여 지도를 수정하였다.

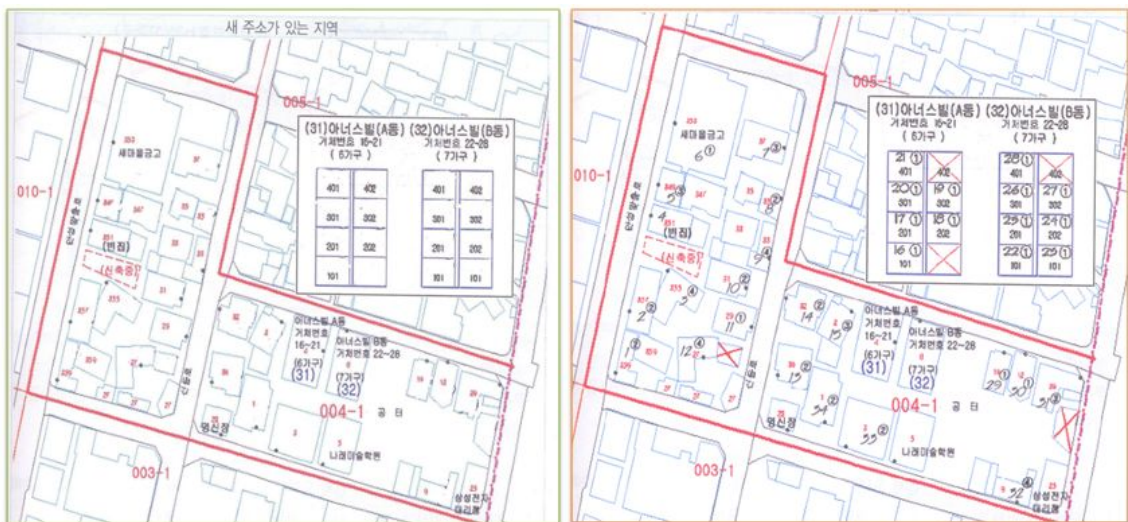
3. 준비조사와 조사구 확정

준비조사는 조사구, 조사구 지도, 조사가구를 확정하기 위한 조사로 본조사 직전에 모든 조사원이 조사현장에 나가 실시한다. 지난 2005년 조사에서

는 전수조사의 경우 10월 29~31일 3일 동안, 표본조사의 경우 30일과 31일 이틀 동안 준비조사가 있었다. 준비조사에서는 조사원이 설정된 조사구에 나가 실제 가구 수를 파악하고, 조사구 지도의 도로와 건물 및 조사구의 경계를 확인하여, 조사구와 조사구 지도를 수정·보안한다.

조사 현장에서 직접 지도 위에 잘못된 경계를 수정하고, 실제와 다른 건물이나 지형지물을 표시하며, 신축 또는 누락된 건물을 그려 넣고, 신축 중인 건물은 점선으로 표시하며, 건물마다 사람이 사는지 여부를 확인하여 거처가 아닌 건물은 "X"표한다(그림 부록7-2). 이와 함께 건물에 거처번호를 부여하고, 거처별로 실제 거주하는 가구 수를 파악하여 조사구 지도에 거처번호와 함께 기입한다.

준비조사과정에서 아파트, 연립주택, 다세대주택 등의 공동주택의 주택 관련 정보를 파악하며, 조사원의 가장 중요한 업무 가운데 하나가 가구명부 작성이다. 가구명부는 본조사에서 조사원이 조사할 대상의 명단이다. 가구명부에는 각 건물의 주소, 지정된 조사구 번호, 공동주택 명칭 등이 포함되어 있다(<표 부록7-3>). 가구명부작성에서 조사원은 조사구 지도에 기입한 거처번호와 가구번호를 기입하는데, 이렇게 작성한 번호는 조사구번호와 함께 가구의 고유한 ID가 된다.



(그림 부록7-2) 조사구지(요)도 사용 방법 예시 (조사 전 → 조사 후)
출처 : 통계청(2004a)

<표 부록7-3> 공동주택 기본정보 수집표(양식) 예시

조사구 번호	공동주택 분류			명칭 및 동·호수	연건평 (전용면적)	총방수			건축 년도	편의 시설수			비 고
	아 파트	연 립	다 세대			방 수	거 실 수	식 당 수		부 엿 수	화 장 실 수	출 입 구 수	
032-A	○			○○○아파트 201동 101호 ~ 2002호까지	26평	3	1	0	1998	1	2	1	
.....													

출처 : 통계청(2004a)

<표 부록7-4> 가구명부 예시

시·도	시·군· 구	읍·면· 동	조사구 번호	조사표 부수	가구명부 매수	조사구 지도 매수	조사원	남조사
○○○시	○구	○○○동	31-1		매중 매	매	조사관리자 총관리자	박인구 김통계

거 처 번 호	가 구 번 호	주소					가구주 성명	가구 원수		가구 종류				거처 종류				농림어가 여부			비 고
		법정 읍면동	동리	번지	공동주택			남	여	외 국 인 가 구	단 독 주 택	아 파 트	연 립 주 택	다 세 대 주 택	비 거 주 용 건 물	주 택 이 외 거 처	농 가	임 가	어 가		
					명칭	동· 호수															
001	001	○○동	3통	121-1	○○ 빌라	A동 101호															
002	001	○○동	3통	122																	
	002	○○동	3통	122																	
003	001	"	"	123																	

출처 : 통계청(2004a)

제2절. GIS 기반의 조사구 설정 및 관리·활용

1. GIS를 기반의 조사구 설정 및 자료의 DB 구축

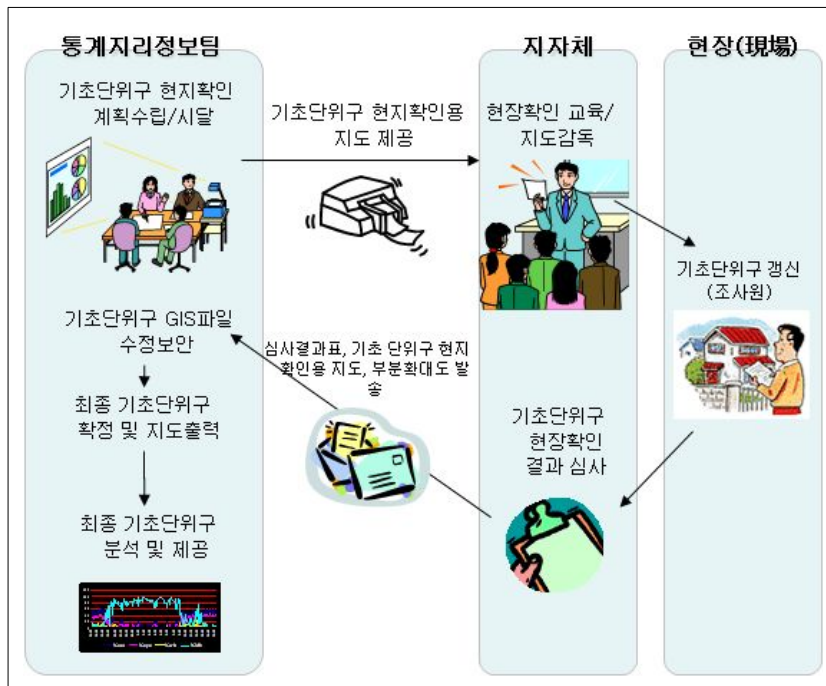
앞에서 보았듯이 조사구 설정작업은 명실공이 발로 뛰는 작업이다. 그나마 지난 10년에 걸쳐 수치지도로 조사구 지도를 만들어 들고 나갈 수 있었던 것은 현장에서 큰 도움이 되고 있다. 조사구 지도를 만들기 위해 통계청은 각 기관들로부터 협조를 받아 구축한 1:1,000 수치지형도를 만든다. 다음은 이 지도를 바탕으로 사전조사단계에서 조사원들이 현장에 나가 거처확인

및 조사구요도, 가구명부를 확인하는 작업을 수행하며, 이를 다시 내부적으로 60가구 내외가 되도록 조정하고 있다.

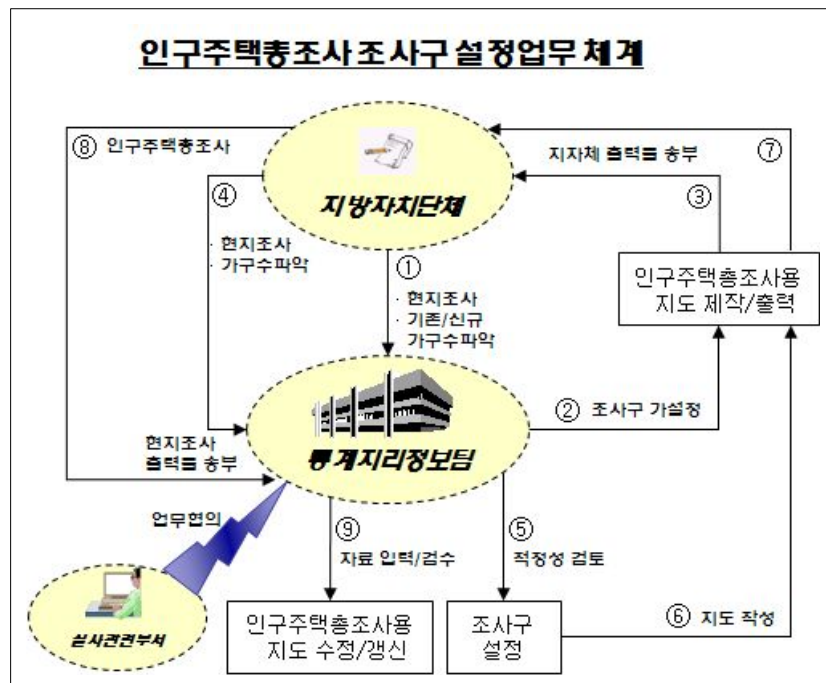
문제는 인구주택총조사를 위한 조사구 설정업무는 인구주택총조사의 특성상 5년마다 이루어지고 있는데, 한번 활용된 조사구 설정관련 자료 (거처, 조사구 요도, 가구명부자료등)는 인구조사과에서 재활용되는 것이 아니라 그냥 사장되고 있다는 사실이다. 물론 통계지리정보팀에서 이 자료를 DB화 하여 조사업무에 활용할 수 있도록 할 예정이지만 통계지리정보팀에서 구축한 자료를 바탕으로 2010년 사전준비단계에서 조사구를 설정하고, 현장에서 이를 보완하도록 한다면 인구주택총조사의 조사구 설정업무는 훨씬 효율화될 수 있을 것으로 판단된다.

특히 인구주택총조사는 5년마다 수행되지만 통계청에서 매해마다 수행하는 사업체기초통계조사에서도 인구주택총조사와 유사하게 조사구를 설정하는 업무가 수행되고 있으며, 매해 조사구를 설정하면서 기초단위구³⁾ 자료도 갱신되고 있기 때문에 이들 자료를 활용하면 현장에서 확인해야 하는 업무량도 많이 감소할 수 있을 것이다. 이러한 상황은 특히 인구주택총조사에서 지방자치단체의 협조를 받을 수 있는 지원체계가 갈수록 미흡해지는 상황에서 기 구축된 자료를 정보화하고 이를 체계적으로 관리할 수 있는 기틀을 마련하는 것은 매우 필요한 상황이다.

3) 기초단위구에 관해서는 다음 장 참조



(그림 부록7-3) 통계청의 기초단위구 관리 업무현황
출처 : 통계청, 2006c



(그림 부록7-4) 통계청의 조사구 관리 업무현황
출처 : 통계청, 2006c

2. GIS 기반으로 다른 업무와 연계 및 결과 활용

한편, 조사구설정과정 가운데 사전준비단계에서 조사원들은 통계청이 구축한 조사구요도를 들고나가 현장에서 건물이나 도로 등 달라진 지형지물을 표시하고, 국민들이 살고 있는 거처에 대한 파악과 함께 거처별 가구수를 확인하는 작업을 수행한다. 이 과정 중에는 현장자료 즉 현장의 변동사항을 반영한 수치지형도가 필요하며 거처의 확인을 위해 건축물에 대한 행정자료의 확보는 조사구 설정업무에 큰 도움이 될 수 있다.

현재는 매해 실시되는 사업체기초통계조사와 기초단위구 갱신에서 이들 지형지물을 갱신하는 일을 수행하고 있으나, 현재 행정자치부를 중심으로 추진중에 있는 새주소 사업 데이터베이스는 데이터의 정확도는 상대적으로 떨어질 수 있지만 건물 및 도로의 변동사항을 체계적으로 유지관리될 수 있는 기틀이 만들어져있는 상황이므로, 이들 자료를 협조 받아 활용하는 방안을 모색하는 것이 필요하다. 건축물의 정보와 관련하여서도 앞의 여러 논문에서 건축물대장의 누락이나 정확도의 한계에 대해 지적을 하고 있으며, 그러한 지적도 사실이기는 하나 이는 과거자료까지 포함을 하기 때문에 발생하는 문제이며, 지속적으로 갱신되는 데이터들, 즉 예를 들면 2005년 이후 갱신된 건축물대장정보들 만을 협조받아 활용한다든지 할 때에는 상대적으로 정확하기 때문에 이러한 행정자료를 활용할 수 있는 방안을 적극적으로 모색하는 것이 필요하다.

이렇게 관련기관에서 보유하고 있거나 관리하고 있는 자료의 협조를 받음으로써 조사구 설정과 관련된 업무에 있어 통계청은 통계청본연의 업무에 충실할 수 있으며, 조사구설정을 위해 지형지물의 변동사항을 유지·관리해야 한다든가 달라진 건축물 정보를 갱신해야 하는 업무들에 대해서는 부담을 줄일 수 있을 것으로 판단된다.

참고 : 미국의 GIS 기반 조사관리

미국은 조사에서 지리정보를 활용하는데 적극적이다. 1990년 인구센서스에서는 TIGER (Topological Integrated Geographic Encoding and Referencing)를 개발하여 활용하였다. TIGER는 지형지도, 도로, 강, 철도 등의 도로지도, 건물과 시설물 지도에 지역경계선, 센서스 블록 등의 조사 집계 단위에 관한 정보가 함께 포함된 지리정보체계(Geographical Information System)다.

미국의 지리정보 활용은 1960년대부터 시작되었다. 당시는 조사응답자의 위치를 기록하여 관리하는 데 종이로 된 지도를 사용하였다. 1970년대에는 센서스를 지리코드로 처리하고 참조파일을 자동으로 편집할 수 있는 GBF(Geographic Base File)/DIME(Dual Independent Map Encoding)를 개발하여 사용하였다. 1980년대에는 조사표 발송 및 회수에 지리코드를 부여하여 사용하였다. 1990년대에는 TIGER 파일을 이용하여 센서스블락 단위 집계할 수 있도록 하였다. 2000년대에는 GIS를 기반으로 조사관리 및 운영을 자동화하였다. 또한 GPS를 바탕으로 수치좌표상에 가구별로 응답자의 지역코드를 부여하였으며, 우편조사표가 회수되지 않을 경우 조사원이 PDA에 탑재된 전자지도로 현지에서 무응답 가구의 위치를 파악하여 조사한다.

제2장. 기초단위구 개요 및 활용

제1절. 기초단위구의 개념 및 획정 기준

기초단위구는 도로, 하천, 철도, 산능선 등과 같은 준항구적인 명확한 지형지물을 이용하여 지도상에 구획한 최소 단위구역으로 2000년 인구주택총조사에서 처음으로 도입되었다. 기초단위구의 구획은 먼저 읍면동별로 가구수에 관계없이 간선도로, 주요 하천, 산능선 등과 같은 명확한 지형지물을 이용하여 크게 대구역을 구획하고 대구역 내에서 다시 소구역을 구획하는 방식으로 이루어진다. 기초단위구 구획의 구체적인 기준과 방법은 다음과 같다(<표 부록7-5>).

가구수에 관계없이 읍면동 내에서 먼저 대도로, 대하천, 산맥 등과 같은 주요 지형지물을 이용하여 대구역을 구획하고, 그 다음 명확한 지형지물을 이용하여 소구역을 구획한다.

- 시가지 조성으로 인하여 블록화된 지역은 가구수에 관계없이 블록 단위로 구획하고 아파트지역은 건물 동 단위로 구획
- 시가지 주변지역이나 농촌지역은 가급적 자연부락 단위로 구획하였는데, 명확한 지형지물이 없는 경우에는 통·리 경계를 이용하여 구획
- 공단 등 사업체 밀집지역은 공단내 도로, 하천 등을 기준으로 구획
- 시가지 변두리지역 또는 농촌·산간지역은 면적이 다소 크더라도 가급적 자연부락이 분할되지 않도록 구획. 명확한 지형지물 이용이 곤란한 경우, 통리 경계를 활용
- 도서지역은 섬 단위로 구획하되, 자연부락이 2개 이상인 경우 분할해서 기초단위구 획정
- 개발이 이루어지고 있는 지역은 전체를 하나의 기초단위구로 설정하며 개발이 완료된 후에 구획기준에 의해 분할

<표 부록7-5> 대구역과 기초단위구(소구역) 구획 기준

구분	대구역	기초단위구(소구역)
도로	<ul style="list-style-type: none"> - 교통량 해소를 목적으로 건설한 고속국도, 지방도·국도 등 - 도시지역은 도시계획법에 의한 광로, 대로, 중로 등 • 광로(1류: 70m 이상, 2류: 50~70m, 3류: 40~50m 미만) • 대로(1류: 35~40m, 2류: 30~35m, 3류: 25~30m 미만) • 중로(1류: 20~25m, 2류: 15~20m, 3류: 12~15m 미만) 	<ul style="list-style-type: none"> - 도시지역: 도시계획법에 의한 소로 등 • 소로(1류: 10~12m, 2류: 8~10m, 3류: 3~8m 미만) - 기타지역: 도로폭 3m 이상 일반도로 등
하천	<ul style="list-style-type: none"> - 하천법에 의해 하천등급이 지방하천 2급 이상 • 국가하천:한강, 낙동강, 금강, 섬진강, 영산강 권역의 약 62개 하천 • 지방하천 1급: 한강권역(청계천 등) 등 5개 하천권역의 55개 하천 • 지방하천 2급: 한강권역(양재천, 세곡천 등) 등 5개 하천권역의 3,799개 하천 - 하천등급이 지방하천 2급 이하 (준용 하천 등)이며 하천폭이 10m 이상 	<ul style="list-style-type: none"> - 하천폭이 10m 미만 - 명확하게 식별할 수 있는 하천 줄기(상하수로 등)
산	<ul style="list-style-type: none"> - 산능선이 길게 형성(산맥)되어 생활권이 다른 경우 	<ul style="list-style-type: none"> - 읍면동내의 통·리 경계가 산능선을 따라 형성되어있는 경우 활용
철도	해당없음	<ul style="list-style-type: none"> - 국철 또는 사철: 준영구적인 철로만 해당

출처 : 통계청(2003a)

제2절. 기초단위구 지정 및 관리 현황

기초단위구는 2000년 인구주택총조사부터 사용되었는데, 2003년 10월 현재 전국 234개 시군구, 3,552개 읍면동에 걸쳐 총 321,174개가 설정되어 있다. 최초 설정 이후 지형지물 및 행정구역 경계 변동 등 지역변화를 반영하기 위해 2001년, 2002년, 2003년, 2005년 등 총 4회에 걸쳐 기초단위구에 대

한 수정작업이 진행되었으며, 이를 통해 도로 신설 및 폐지, 아파트 신축 및 철거, 재개발, 구획정리 등 지역변동과 대구역 및 소구역의 특성 변동 사항을 반영하였다.

기초단위구수는 2002년 316,857개에서 2003년 321,174개로 약 1.4%가 증가하였는데, 변동내용을 전국 단위에서 살펴보면, 도로신설, 아파트신축, 재개발 등으로 인해 기초단위구가 분할된 경우가 5,802건이며, 도로폐지, 아파트철거, 재개발 및 구획정리 등으로 인해 기초단위구가 통합된 경우가 1,485건으로, 기초단위구의 변동이 이루어진 경우는 2002년 전체 기초단위구 기준으로 약 2.3%에 해당한다(<표 부록7-6> 참조).

<표 부록7-6> 시도별 기초단위구 현황 및 변동내역

구 분	시군구 수	읍면동 수	행정구역 면적(km ²)	기초단위구수		변동률* (%)	변동		
				2002년	2003년		소계	분할	통합
전국	234	3,552	100,917	316,857	321,174	2.3	7,287	5,802	1,485
서울특별시	25	522	607	59,781	60,061	1.9	1,152	716	436
부산광역시	16	222	777	23,456	23,729	2.3	539	406	133
대구광역시	8	143	881	15,712	15,926	2.3	364	289	75
인천광역시	10	139	1,084	15,155	15,247	1.1	170	131	39
광주광역시	5	90	499	8,663	8,959	4.1	352	324	28
대전광역시	5	79	540	9,534	9,782	3.1	294	271	23
울산광역시	5	58	1,068	7,185	7,284	1.6	117	108	9
경 기 도	31	506	10,282	58,606	60,123	3.7	2,183	1,850	333
강 원 도	18	188	16,956	12,039	12,095	1.0	118	87	31
충청북도	12	153	7,433	12,115	12,378	3.4	409	336	73
충청남도	16	209	8,781	13,360	13,480	1.6	218	169	49
전라북도	14	251	8,007	16,383	16,421	0.9	140	89	51
전라남도	22	298	12,461	16,637	16,677	1.0	164	102	62
경상북도	23	337	19,055	21,400	21,579	0.9	193	186	7
경상남도	20	314	10,597	22,068	22,469	3.0	673	537	136
제 주 도	4	43	1,888	4,763	4,964	4.2	201	201	-

* 변동률은 2002년 기초단위구수를 기준으로 계산한 값

출처 : 통계청 내부자료

제3절. 기초단위구 번호부여 체계

기초단위구에 부여되는 번호체계는 총 8자리의 숫자로 이루어지는데, 앞에서부터 차례대로 대구역번호, 소구역번호, 분할번호, 특성번호가 각각 2자리씩 부여된다.



(그림 부록7-5) 우리나라 기초단위구 일련번호 구성 체계

대구역 번호와 소구역 번호는 각각 읍면동, 대구역 내에서 좌측 또는 우측 상단부터 순차적으로 연결하여 부여하게 된다. 분할번호는 기초단위구 획정 후 택지개발이나 지형지물의 변화로 인해 해당 기초단위구의 분할이 일어날 경우에 부여되는데, 기초단위구 설정시는 “00”으로 부여하고 기초단위구 설정 후 택지개발 또는 지형지물의 변화로 분할될 경우 부여하게 된다. 마지막으로 특성번호는 기초단위구내 지형지물과 면적기준으로 대표적인 특성을 분류기준에 의해 대분류와 소분류로 부여하는 것으로, 해당 기초단위구의 토지이용특성을 나타낸다.

<표 부록7-7> 기초단위구 특성번호 분류 기준

대분류	분류기준	소분류	분류기준
① 주거 지역	주택 70%이상	① 단독주택지역	단독주택 70% 이상
		② 아파트지역	아파트 80% 이상
		③ 공동주택지역	공동주택(연립, 빌라, 맨션 등) 80% 이상
		④ 주택혼합지역	주택종류가 70%이상 혼합된 주택지역
② 준주거 지역	주택 30~70%	① 단독주택, 상가 복합지역	단독주택이 30~70%이고, 상가와 혼합
		② 공동주택, 상가 복합지역	공동주택이 30~70%이고, 상가와 혼합
		③ 단독주택, 공장 복합지역	단독주택이 30~70%이고, 공장과 혼합
		④ 공동주택, 공장 복합지역	공동주택이 30~70%이고, 공장과 혼합
		⑤ 기타복합지역	주택이 30~70%이고, 공장 및 상가와 혼합
③ 비주거 지역	주택 30%미만	① 상가밀집지역	주택이 30% 미만이고, 상가 50%이상
		② 공공시설지역	주택이 30% 미만이고, 공공시설 50% 이상
		③ 문화시설지역	주택이 30% 미만이고, 문화시설 50% 이상
		④ 공장밀집지역	주택이 30% 미만이고, 공장 50%이상
		⑤ 상가, 공장밀집 지역	주택이 30% 미만이고, 상가와 공장이 각각 50% 이상
		⑥ 기타 복합지역	주택이 30% 미만이고, 상가 50%이상
④ 농어촌, 산간지역	농경지, 산, 하천, 해안 등이 70% 이상	① 평야지역	평야지대로 농경지가 70% 이상
		② 준평야지역	농경지가 30~70% 미만
		③ 산간지역	산, 하천 등이 30~70%
		④ 해안지역	해안선(바다)과 접해있는 지역
⑤ 섬지역	사면이 바다로 둘러싸인지역	① 섬지역	시군청 이상의 행정기관 소재지 섬은 제외

출처 : 통계청(2003a)

<표 부록7-8> 일부지역의 기초단위구별 가구수 (2000년 기준)

구 분		서울시 종로구			서울시 강남구			전주시		
		행정동	대구역	기초구	행정동	대구역	기초구	행정동	대구역	기초구
개수		19	45	1,506	26	84	3,534	40	132	4,923
가구 (호)	전체	56,302	-	-	170,929	-	-	182,556	-	-
	평균	2,963	1,251	40	6,574	2,164	54	4,564	1,383	37
	최소	1,428	41	1	2,035	1	1	1,281	7	0
	최대	4,973	4,154	435	11,464	5,812	472	12,567	5,537	355
구 분		용인시			평창군			태안군		
		행정동	대구역	기초구	행정동	대구역	기초구	행정동	대구역	기초구
개수		14	45	2,601	8	9	324	8	10	428
가구 (호)	전체	108,949	-	-	14,160	-	-	21,615	-	-
	평균	7,782	2,421	42	2,047	1,820	44	2,702	2,162	51
	최소	1,941	287	0	1,009	1,009	0	858	858	0
	최대	32,705	7,513	306	3,975	3,112	128	8,525	3,495	253

출처 : 강영옥 외 (2007)

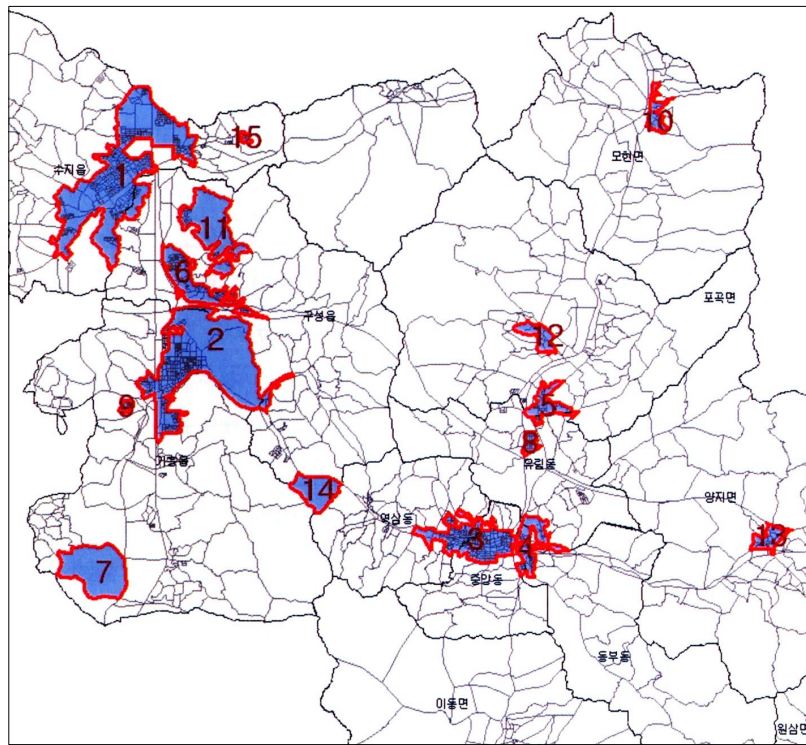
제4절. 기초단위구의 활용

1. 도시화지역 설정

1995년부터 도시지역과 농촌지역이 통합된 도농복합 형태의 시가 나타나면서, 종전에 행정구역을 토대로 도시지역으로 간주되어온 5만 이상의 시지역 (혹은 2만 이상의 읍지역 포함)의 성격이 모호해졌다. 외국에서 일반적인 행정구역과는 별도로 인구밀도 및 토지이용기준을 활용하여 도시지역을 정의해서 확정하고 있으며, 도시화지역 설정도 이러한 필요성에 의해 시작되었다. 2004년과 2006년 수행된 연구에서 (통계청, 2004c, 2006b) 도시화지역은 통계청의 기초단위구를 도시화지역 획정의 기본단위로 활용하였으며, 첫째, 인구밀도 3,000명/Km² 이상을 하거나 지목으로 정의된 도시적 토지이용 면적 비율이 1/2을 초과하는 기초단위구를 도시화지역 획정대상에 포함시키고, 둘째, 이상 두 기준 중 최소한 하나를 만족하면서 공간상에 연결해 있는 복

수의 기초단위구와 그러한 기초단위구들로 둘러싸인 내부의 기초단위구들로 구성된 군집을 도시화지역 후보지로 추출하였다. 이 후 이들 지역 중 인구 3,000명 이상인 지역을 도시화지역으로 최종확정하는 기준을 설정하였다.

이러한 도시화지역 설정기준을 서울특별시와 경기도의 시·구를 대상으로 실제 적용하였으며, 도시화지역 설정을 위한 시스템을 개발하여 향후 도시화지역의 설정과 유지, 관리에 활용하도록 하였다. 현재 통계청에서는 전국에 대한 도시화지역 설정작업에 대한 용역사업이 발주된 상태이며, 2008년 초까지 전국에 걸친 도시화지역 획정이 완료될 예정이다.



(그림 부록7-6) 도시화지역 획정 사례-용인시
출처 : 통계청(2006b)

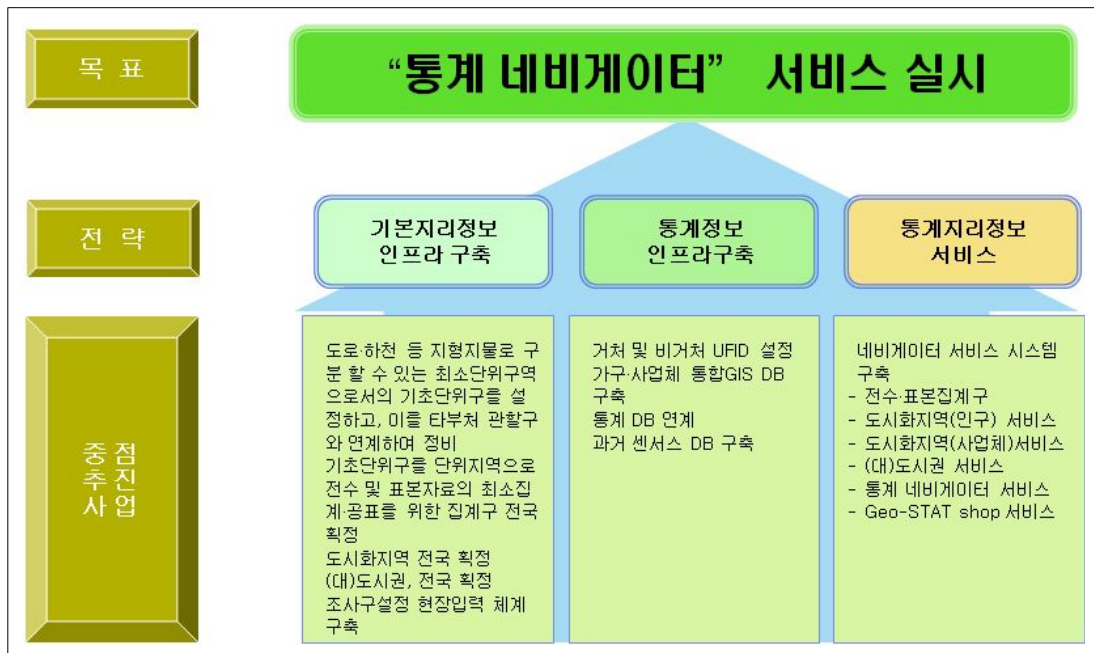
2. 도시권설정

도시권이란 교외화의 진전으로 중심도시를 중심으로 주변지역과의 기능적 연계가 이루어지는 권역으로 우리나라에서는 계획수립을 위해 도시계획권역을 설정한 경우는 많으나, 도시권에 대한 표준적인 획정기준 및 공식적인

도시권 회정이 이루어진 적은 없는 실정이다. 그러나 미국, 캐나다, 프랑스, 일본등에서는 통계청 등에서 도시권을 정의하고 주기적으로 확정하여 정책적 수요에 대응하고 있다. 통계청에서도 이러한 수요를 파악하고 이에 대응하기 위해 2007년에 ‘도시권 확정’과 관련된 연구용역을 발주한 상황이며 현재 과업이 진행중이고, 이 사업은 2007년 12월에 완료될 예정이다.

3. 기타

통계지리정보과에서는 이 외에도 통계정보 인프라로서 지리정보와 인구 및 사업체 통합 DB를 연계하여 서비스가 가능하도록 거처/비거처별 좌표정보를 구축하는 거처 및 비거처의 UFID구축사업, 통계지리정보 서비스로서 이용자가 전국의 어느 지점이나 행정구역, 주소, 상호, 우편번호 등을 이용하여 지도를 찾아내고 지도에서 세분화된 센서스 자료 등의 통계를 검색하여 볼 수 있는 서비스인 통계네비게이터 서비스 사업등을 실시할 예정이다.



(그림 부록7-7) 통계지리정보팀 향후 업무계획

출처 : 통계청(2006c)

제3장. 모바일 GIS 개요 및 활용사례⁴⁾

1. 모바일 GIS를 이용한 현장업무의 효율성 확보

인구주택총조사업무에 있어서는 현장에서 이루어지는 업무들이 많이 있다. 사전준비단계에서 조사구 설정을 위해 통계청이 보유하고 있는 수치지형도를 현장조사 때 갱신하며, 거처의 확인, 조사구역도 확인, 거처별 가구의 확인 등의 작업이 이루어진다. 이러한 작업은 한 번에 끝나는 것이 아니며 여러 차례의 확인과정을 거치게 된다. 조사구 설정이 끝나고 실제 조사는 15일 동안에 이루어지게 된다. 이러한 과정 중에 포함되어 있는 현장조사 작업은 통계청에서 현장조사를 위해 필요한 도면출력, 조사자료 준비, 현장에서 확인을 통한 조사구역도에 변동사항 기입, 가구명부의 작성, 그리고 현장에서 기입된 자료들을 다시 내부적으로 데이터베이스화하는 단계를 거치게 된다. 그러나 모바일 GIS로 현장업무를 지원하는 시스템을 만든다면 조사원들은 PDA등을 이용하여 도면을 출력할 필요도 없고, 자신의 조사지역에 대해 외부작업이나 내부작업이 필요 없이 현장에서 PDA에 수정정보를 바로 입력하여 업무도 효율화하고, 자료처리에 소요되는 시간도 경감할 수 있다.

통계청 전체적으로 본다면 조사를 위한 현장작업이 인구주택총조사 뿐만 아니라 사업체기초통계조사를 비롯한 다양한 조사업무가 있으며, 조사업무에 있어 위치정보의 활용이 점차 많아지고 의존적이 될 수 밖에 없음을 감안할 때 현장업무를 효율적으로 지원해줄 수 있는 모바일 GIS의 도입은 필연적인 것으로 판단된다.

지금까지 인구주택총조사에서 GIS의 활용을 조사업무의 효율성확보를 위한 방안에 초점을 두어 살펴보았다. GIS의 활용은 업무의 효율성 확보 뿐만 아니라 조사된 인구주택총조사의 조사결과를 좀 더 빠르게 국민들에게 서비스하고, 좀 더 다양한 공간단위로, 수요자의 요구사항에 맞추어 서비스할 수 있는 기반이 될 수 있다. 그러나 이러한 다양한 서비스체계 구축과 관련하여서는 현재 통계지리정보팀에서 다양한 계획을 수립하여 진행 중이므로, 이러한 노력들이 차질 없이 진행된다면 GIS가 수요자요구에 맞는 서비스 제공이

4) 강영욱(2006)에서 발췌

라는 소기의 목적을 달성하는 데에도 효율적으로 활용될 수 있을 것으로 판단한다.

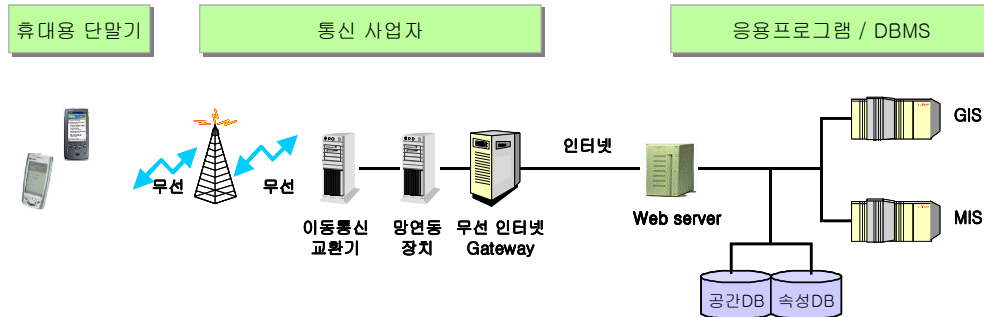
2. 모바일 GIS의 개요

최근의 GIS는 정보통신기술의 급속한 발전에 의하여 데스크탑 단말기가 아닌 현장(Field)에서 업무를 처리할 수 있는 모바일 단말기를 이용한 모바일 GIS로 추세가 확산되고 있다. 모바일 GIS는 하드웨어 및 소프트웨어, GPS, GIS, 무선통신기술이 통합된 정보기술의 집합체로서 여러 분야에 다양한 콘텐츠를 제공하고 있다. 모바일 GIS 시스템의 가장 큰 특징은 이동성, 현장성, 적시성 등으로 공간 및 속성과 관련되는 각종 자료를 현장에서 확인하고, 조사된 자료를 바로 입력할 수 있다는 점이다.

모바일 GIS는 모바일 기기에 지도를 포함한 다양한 GIS응용프로그램을 장착하여, 장소 및 시간의 제약 없이 위치정보 및 지리정보 획득이 가능한 시스템으로 정의될 수 있다. 모바일 GIS는 초기에 모바일용 컴퓨터에 한정되었으나, 현재는 기존의 GIS시스템을 휴대용 단말기로 이식시켜 놓은 것 이외에 모바일 컴퓨팅 환경에 적용할 수 있는 새로운 시스템이라 정의할 수 있다. 최근에는 PDA, 휴대폰, 스마트 폰과 같은 휴대용 단말기의 개발과 대용량의 자료를 전송할 수 있는 무선 인터넷 기술의 발전으로 휴대성과 와이어리스를 모두 충족하는 온라인 모바일 GIS를 일반적인 모바일 GIS 개념으로 사용하고 있으나 오프라인 모바일 GIS 또한 모바일 GIS 개념에 포함되고 있다.

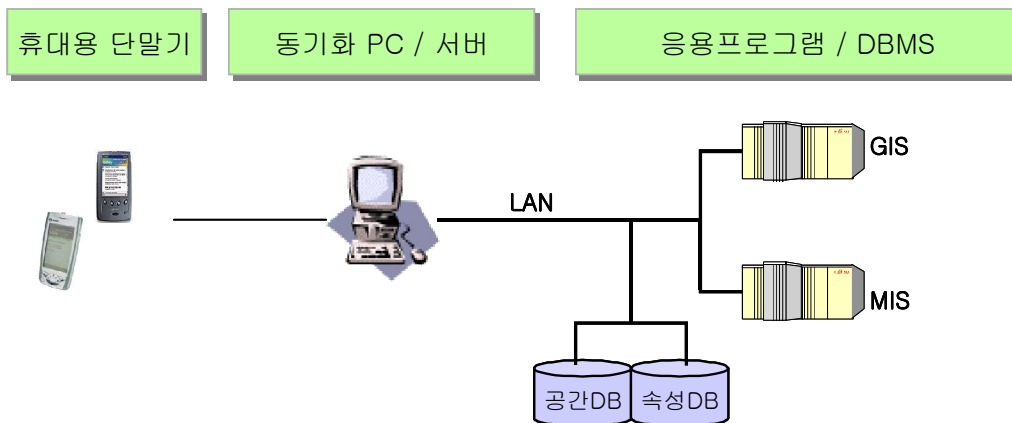
무선네트워크 기반의 온라인 모바일 GIS 구조는 휴대용 단말기, 무선게이트웨이, 무선통신망, 모바일 GIS 응용시스템으로 구성된다. 모바일 GIS 사용자는 휴대용 단말기를 가지고 통신사업자가 제공하는 무선 인터넷망을 통해 원격지의 GIS응용시스템에 접근하며, 무선 인터넷망은 GIS 응용시스템 쪽 웹서버를 통해 GIS 응용시스템과 통신사업자 사이를 유선 인터넷으로 연결하고, 통신사업자 측에서 제공하는 무선게이트웨이, 네트워크 연동장치, 이동

통신 교환기와 각 기지국을 통해 단말기와 연결되는 구조를 나타낸다((그림 부록7-8)).



(그림 부록7-8) 온라인 모바일 GIS 구조
출처 : 서울시정개발연구원(2002)

오프라인 모바일 GIS 구조는 휴대용 단말기, 동기화 서버, 모바일 GIS 응용시스템으로 구성되는데, 사용자는 휴대용 단말기에 설치된 응용프로그램을 사용하기 위해 GIS 응용시스템에 저장된 공간자료를 다운로드 받거나, 휴대용 단말기를 이용하여 현장에서 수집한 정보를 GIS 응용시스템에 올려 업무에 필요한 작업을 수행하며((그림 부록7-9)), 이 작업은 GIS 응용프로그램과 인터넷이나 LAN으로 연결된 동기화 서버를 사용하게 된다.



(그림 부록7-9) 오프라인 모바일 GIS 구조
출처 : 서울시정개발연구원(2002)

모바일 GIS는 휴대가 간편하고 값싼 단말기를 사용하여 GIS 시스템에 접

근 가능하고 실시간으로 현장의 데이터를 입력, 편집, 수정할 수 있는 뛰어난 현장성을 갖고 있어 현재는 현장 업무 중심의 GIS 시스템(Field GIS), 지도제작, 위치기반서비스 분야 등에서 공간자료의 분석, 제공, 수집, 편집, 갱신 작업 등에 적용되고 있다.

<표 부록7-9> 모바일 GIS 응용구분

응용분야	서비스 내용
Field GIS	<ul style="list-style-type: none"> · 휴대용 단말기와 무선 통신망을 이용한 자료 수집과 무선 통신망을 이용한 자료의 편집, 갱신 · 휴대용 단말기를 통해 시설물 정보를 확인하여 현장에서 정보의 신뢰성을 검증 · 업무 내용에 따른 현장 정보 제공과 지역 분석 · 교통, 조세 등의 현장 단속 업무에서 현장에서 단속내용을 실시간으로 전송
Mapping	<ul style="list-style-type: none"> · 휴대용 단말기와 GPS 수신기를 사용하는 지도 제작 · 휴대용 단말기를 통해 인구, 관광, 문화 등 다양한 지역정보를 가시화
LBS	<ul style="list-style-type: none"> · 주변지역 정보 제공 · 특정 지역에 도달하기 위한 경로 찾기, 위치추적 · 휴대전화 사용 위치에 따른 차별화된 사용요금 부과 · 긴급구난 서비스에서 사고자의 위치를 파악

출처 : 서울시정개발연구원(2002)

3. 모바일 GIS 적용 사례

가. 국내 모바일 시스템 구축 사례

현재까지 공공부문에서 구축된 모바일 시스템의 응용분야는 단속 업무, 시설물관리업무, 조사업무, 소방/방재/치안업무가 주류를 이루며, 대민서비스 분야에서 활용되고 있는 사례는 민원정보, 생활정보, 관광정보 등으로 구분할 수 있다. 민간분야에서는 Web GIS기술을 이용하여 인터넷으로 지도서비스를 하고 있는 많은 GIS 벤처회사에서 m-GIS 기술과 GPS 기술을 응용하여 이동 중에 지도서비스를 제공받는 위치확인 시스템을 많이 활용하고 있는 상황이다.

단속 업무에서 활용되는 모바일 시스템은 무선통신기기를 이용하여 주차 단속, 환경오염 상태등을 단속하고 단속정보를 곧바로 중앙서버에 전송하여 요금을 징수하는 서비스를 제공한다. 특히, 주정차단속업무에서의 활용은 서울시의 각 구청 및 다른 시도의 자치구에서 활발히 적용되고 있는데, 현장에

서 즉시 입력함으로써 민원에 실시간으로 응대할 수 있으며, 사무실에 들어와서 입력하는 이중 업무를 감소시키는 효과가 있다.

시설물 관리업무에서 활용되는 모바일 시스템은 각종 시설물에 대한 관리 업무를 수행하는 현장 근무자와의 업무 연락, 지시사항을 PDA를 통하여 전달함으로써 현장에서 처리되는 내용이 직접 저장되고 시설물에 대한 유지관리, 고장접수, 민원 관리 등의 정보를 제공한다.

조사업무에서 활용되는 모바일 시스템은 현장에서 수집된 조사내용을 실내에서 전산 입력하는 대신 현장에서 PDA 등의 이동형 단말기를 통하여 직접 입력하도록 하였다. 이를 통해 조사표에 수기로 작성하던 업무와 실내에서 전산으로 입력하던 이중 업무를 단축할 수 있으며, 토지나 건물 조사업무에서 기존의 아날로그 형태의 지도를 이용하는 경우보다 도면출력과 위치 확인의 소요시간과 비용을 대폭 절감할 수 있다⁵⁾.

소방/방재/치안 등의 업무에서는 휴대폰을 가진 구조 요청자가 보낸 구조 신호를 인공위성과 LBS 시스템을 통하여 요청자의 위치정보로 변환하여 별도의 위치보고 없이 해당자의 위치정보를 파악할 수 있어 긴급출동 등에 활용된다. 또한, 휴대폰, PDA 등의 단말기에 실시간으로 업데이트 되는 재난 속도 등 정보를 제공한다.

대시민서비스 분야에서는 민원, 생활, 관광 등의 정보를 실시간으로 제공하기 위해 모바일시스템을 도입하고 있다. 민원정보는 개인 휴대단말기(PDA) 등 모바일기기를 통하여 문자나 음성으로 민원안내, 신청, 처리결과 통보 등의 각종 민원서비스를 제공하며, 이 서비스를 통해 민원 업무처리까지연되는 상황을 막고 민원인에게 각종 행정업무 편의를 제공하고 있다. 생활정보의 하나인 버스정보는 버스관리시스템(Bus Management System : BMS)에서 관리하고 있는 버스운행시각과 도로상황 등의 정보를 모바일기기를 통해 시민들에게 실시간으로 제공함으로써 시민들의 편의성을 증대하고, 대중교통 활성화를 기할 수 있다. 그 외 무선관광정보시스템을 통해 관광지의 관광정보 제공 및 추천 코스 안내뿐만 아니라 관광에 필요한 각종 교통 및 도로정보, 교통편의 기능을 제공하며, 숙박시설 및 식당 등 업소에 대한 주제별·지역별 추천 및 안내를 통해 관광산업 활성화에 기여할 수 있다.

5) ITtimes, 2006.3. 김포신도시 토지현장조사에 개인휴대용 정보단말기 도입

<표 부록7-10> 위치기반서비스(LBS)와 차량항법시스템(ITS)

분류	구축 시스템	제공되는 정보 및 업무	시행지역	
단속업무	PDA 주정차관리시스템	· 주정차단속업무, 단속사항 조회업무	서울특별시, 인천광역시 외	
	차량단속업무시스템	· 차량단속업무, 자동차체체납관리업무	광주광역시/군산시	
	지방체납시스템	· 체납차량조회, 영치, 번호판영치번호판봉인등의 단속업무	경기도	
시설물 관리업무	이동행정서비스 지원시스템	· 상하수도 점검업무, 무선통신을 이용하여 상수도 서버에 접속	대전광역시	
조사업무	농수산물가격조사	· 농수산물시장의 시세를 PDA와 무선네트워크를 통해 실시간으로 입출력	농수산물유통공사	
	소비자물가조사	· 종이조사표 대신 PDA에 직접 입력 후 통계청으로 전송	통계청	
	공시지가조사	· 기구축된 토지특성자료 조회 및 변동사항 PDA에 직접 입력	성동구, 노원구, 울산시등	
소방/방재/치안업무	e119	· 구조요청자가 보낸 구조신호를 통해 위치를 파악하여 긴급출동	행정자치부	
	무선조회	· 경찰전산망에 접속하여 신상, 차적, 운전면허번호등의 기록조회	경찰청	
대시민 정보 제공	민원 정보	시정정보안내시스템	· 민원업무처리통보	서울특별시
		민원편의정보시스템	· 여권발급확인, 공지사항 안내서비스	충청북도
		휴대폰 문자메시지서비스	· 민원처리결과 통보, 행정업무 연락	부산광역시
		모바일 특허민원서비스	· 특허 관련 민원의 처리결과, 관련 정보를 민원인의 휴대폰에 단문전송	특허청
	생활 정보	버스정보	· BMS 또는 BIS에서 모바일 단말기를 통한 교통정보 제공	부천시외
		재난속보	· PDA, 휴대폰등을 통해 재난 정보 제공	소방방재청
	관광 정보	무선관광정보시스템	· 무선관광정보, 사용자 위치정보확인서비스	대구광역시
		경남관광체험 길라잡이 시스템	· 관광정보, 실시간 기상 및 교통정보	경상남도
		사이버제주구축사업	· 관광정보	제주시

출처 : 서울시정개발연구원(2003)

나. 해외 모바일 GIS 구축 사례

해외에서 구축되고 있는 모바일 GIS는 소방/방재/치안, 자료수집, 단속, 행정관리, 행정정보제공, 생활정보제공, 무선 통신 기반을 제공하는 영역으로 구분할 수 있으며, 영역별 활용 내용은 <표 부록7-11>과 같다.

<표 부록7-11> 해외 모바일시스템 활용 현황

구분	내용	사례지역
소방·방재·치안	· New Orleans 에서 총기발포감지 및 경찰에 무선통신할 수 있는 서비스 실시	미국 루이지애나
	· 주정부 가정폭력방지위원회: 민간기업 Verizon Wireless에 의해 무상 제공된 휴대폰을 가정폭력 희생자 가족들에게 지급하여 보호시설이나 경찰에 비상연락할 수 있도록 함	미국 매릴랜드
	· 휴대단말기를 소지하거나 순찰차량에 휴대단말기를 설치하여 실시간으로 중앙 시스템에 접근하여 관련정보를 조회하거나 순찰내용을 전송할 수 있는 시스템 운영 - Boston, Seattle, Pasco, Sacramento 등	미국 매사추세츠, 워싱턴, 워싱턴, 인디애나, 캘리포니아, 텍사스, 펜실바니아 등
	· 산불 진화에 모바일 시스템을 이용하여 산불이 발생하면 헬기에서 산불 진행상황을 수집하여 분석하고, 현장 소방관들은 휴대용 단말기를 통해 분석결과를 전달받아 효과적으로 진화작업을 수행함. · Southern California에서는 소방관리를 위해 화재위치 및 경로 추적을 위한 GIS기반 무선시스템 구축	미국 캘리포니아
	· Houston에서는 제보자의 위치정보를 제공하는 911 무선 긴급네트워크서비스 실시	미국 텍사스
	· 서부 Midland 경찰청에서 청각장애자, 언어장애자 등을 위해 영국 최초로 모바일폰을 이용한 SMS 긴급연락망 서비스 실시	영국 Birmingham
	· 소방대원 및 사고현장 시민의 위치를 확인 할 수 있는 웹기반 시스템 구축. 웹기반 시스템에 무선게이트웨이를 추가함으로써 화재현장에 대한 실시간 정보를 제공함	미국 캘리포니아주 산 마테오 카운티
	· AVL(Automatic Vehicle Location) 시스템을 제설차량에 장착함. 이를 이용하여 제설차량의 위치를 파악하고 신속하고 정확한 작업지령을 하달함	미국 Maryland Howard 카운티
자료수집	· Center for Disease Control and Prevention: 생화학테러관련 사고자료 수집을 위하여 GPS장착 휴대용 단말기 사용 예정	미국 조지아
	· 모기를 매개로 발생하는 질병에 대한 감시체계를 구축(the West Nile Virus system)하고, 질병발생 자료를 수집하기 위해 현장조사원들이 PDA를 활용함	미국 펜실베이니아
	· 2010 센서스 조사시 현장조사데이터 수집시 활용함	미국 통계청
	· 건물안전진단 요원들이 현장에서 건물관련 정보를 조회하고, 검사결과를 데이터베이스에 저장할 수 있는 PDA 시스템 활용하고 있음	캐나다 Edmonton
	· 도로관리과에서 도로표지판 유지관리를 위하여 도입함	미국 Boulder County
· 상수와 하수 기반시설 정보를 수집하는데 이용함	미국 Edgecombe county	
단속	· 주정부에서 주차단속을 위해 휴대용 무선 단말기를 이용하여 티켓 발부 및 본청 데이터베이스에 전송	미국 뉴저지, 텍사스
	· Bellevue시는 시청에서 구축해 놓은 인트라넷, 이메일 시스템, 기타 여러 전산시스템들을 무선시스템과 연결하는 작업을 진행하여 건물안전진단 요원이 휴대단말기를 이용하여 현장에서 건물에 관한 정보를 받거나 진단결과를 전송할 수 있는 시스템 등이 이미 구축되어 있음	미국 워싱턴
	· 주 감사원(State Auditor), 주이용업검사위원회(State Board of Barber Examiners)의 현장 단속 활동과 주내 탁아시설에 대한 허가 및 검사에 휴대단말기 이용	미국 텍사스

(다음 장에 계속)

<표 부록7-11> 해외 모바일시스템 활용 현황

구분	내용	사례지역
행정 관리	· 주 공공기관 근무자들이 휴대용 무선기기를 이용해 이메일 교환할 수 있는 기반을 구축	미국 켄터키
	· Hinds County에서는 미국내 최초로 관할구역에서 무선투표기를 사용, 신속한 투표결과를 산출	미국 미시시피
	· SMS를 이용한 휴대용 모바일 단말기 또는 주요 공공장소에 설치된 키오스크 등을 이용한 전자투표를 시범 실시	영국 9개 지자체
	· 양수장 관리를 위해 PDA를 활용하여 시설물 관리를 하고 있음	캐나다 Richmond
행정 정보 제공	· 무선기기를 통해 데이터 서비스를 받을 수 있는 무선웹포탈 형식의 서비스 제공 -버지니아에서는 미국 지방정부 최초로 휴대용 단말기로 접속할 수 있는 무선웹포탈(www.myVirginia.org/mobilegov) 구축 -뉴욕에서는 무선통신을 이용하여 시민들이 시정부 웹사이트로부터 비상 대비루트 및 대피시설에 대한 정보를 제공받을 수 있도록 함.	미국 뉴욕, 미국 버지니아, 캘리포니아
	· 시청 311 콜센터에 고객센터요구시스템(CSR:Customer Service Request)을 구축하여, 시민들의 민원이 들어오게 되면 시청직원이 외근중 이더라도 무선으로 민원내용을 즉시 전달하여 처리할 수 있도록 하고 있음.	미국 달라스
생활 정보 제공	· 공공버스를 이용하는 사람에게 각 정류장에 버스가 도착할 시간을 알려주는 버스안내시스템을 구축하였음 - Portland, King County 등	미국 워싱턴, 오레곤
	· 각 버스에 GPS 장비를 장착하여 버스의 위치를 파악할 수 있으며, 버스 안내 정보를 이용하려면 무선 인터넷 기능을 갖춘 휴대전화를 사용해야 함.	
	· Salt Lake City: 2002 동계올림픽에서 비상대응 및 차량위치추적을 위해 GPS 및 무선네트워크 사용 · 올림픽 종료 후 긴급서비스차량 추적서비스에 활용)	미국 유타
	· 교통체증 감소를 위하여 타운센터내 빈 주차공간을 운전자들에게 알려주는 SMS 서비스 실시	영국 공통
	· 토지소유권, 등록, 가격 등에 정보를 휴대용 무선단말기를 통하여 제공	영국 공통(토지등기국)
	· 열차 승객들을 위해 모바일 폰에 의한 열차시간표 정보제공	영국 공통
무선 통신 기반 제공	· 시 주차국(Parking Authority)에서는 인터넷 기반의 주차관리시스템을 구축하여 운영하면서 무선인터넷으로 관련 정보를 제공해 주공 있음	캐나다 Toronto
	· 다운타운, 공항, 공원 등에 광대역 무선랜을 설치하여 무료로 무선랜을 이용할 수 있는 기반시설 제공 -Rolling Meadows, Long Beach, Jacksonville, New York 등	미국 뉴욕, 일리노이, 캘리포니아, 플로리다

*출처 : 서울시정개발연구원(2003); 한국건설기술연구원(2002)

<부록 8> 2010년 인구주택총조사 추진일정

1) 사전 준비단계

업 무 명	주 요 내 용	비 고
1. 기본계획 수립 【2007. 1.~7월】	○ 총조사 추진방향 및 일정 수립 [초안 작성 → 청내협의 → 외부전문가의견수렴 → 내부결재 → 시도통계담당관회의 → 시행]	통계청 지방자치단체
2. 조사항목 확정, 조사표 및 결과표 등 설계 【~2009. 12월】	○ 조사항목안 작성→전문가회의→ 조사표 및 결과표 설계→시험조사→시범예행조사→ 확정	통계청, 전문가회의, 지방자치단체
3. 시험조사 【2007~2009년】	○ 실시계획 수립→조사구설정 및용품류 제작·발주→인력동원 및 교육→ 현장조사 및 자료입력	통계청, 지방자치단체
4. 교육준비 【~2009.10월】	○ 교육기법 및 기자재 개발 ○ 조사지침서 및 교안 작성 ○ 교육관련 세부계획 수립	통계청
5. 조사기법 연구·개발 【~2009.10월】	○ 외국인, 노숙자 등 취약지역 조사방법 ○ 인터넷 및 행정자료 활용 조사방안 ○ Imputation 기법 개발	통계청, 외부전문가
6. 홍보전략 수립 【~2009.12월】	○ 홍보기본계획 수립 ○ 홍보 콘텐츠 개발	통계청, 외부전문가
7. 조사용품 개발 【~2009.10월】	○ 용품류 종류 결정→사양 결정 →시범예행조사용 제작→배부계획 수립 및 시행	통계청
8. 조사구 설정 준비조사 【2009. 9.~2010. 2월】	○ 조사구 설정 계획 수립 ○ 사전 가구수 파악 및 현지 확인 ○ 조사구 가설정	통계청, 지방자치단체
9. 실시본부 구성·운영 【2009. 9.~2010. 12월】	○ 실시본부 구성 규정 작성 ○ 전문가 회의체 및 실무자 협의체 구성 ○	통계청 및 실시기관
10. 시범예행조사 【2009. 11.~12월】	○ 총조사 전반적 사항 종합 리허설 ○ 지자체 공무원 참여 및 워크샵 실시	통계청, 지방자치단체, 관련학계 등

2) 본조사 추진단계

업 무 명	주 요 내 용	비 고
1. 종합 시행계획 수립 【 2009.12~2010.3월 】	○ 업무단계별 세부 시행지침 수립·시행 → 관련기관 회의→계획 시행 및 업 무추진	통계청, 지방자치단체
2. 조사구 설정 【 2010. 3~10월 】	○ 가설정 조사구 적합성 확인→표본조 사구 추출→ 조사구 요도 작성→변 동조사구 관리	통계청, 지방자치단체
3. 조사표류 인쇄 및 용품 류 제작·배부 【 2010. 3~10월 】	○ 조사표류 및 용품류 확정→수량산정 및 배부계획 수립→제작·발주→ 배부	통계청, 지방자치단체 지방통계사무소
4. 특별조사구 【 2010. 3~11월 】	○ 관계기관 회의→조사항목 및 조사표 확정→조사추진 및 조사표 수집	통계청, 관련기관
5. 인력동원 【 2010. 7~9월 】	○ 인력동원 세부계획 수립→채용공고 및 등록→채용 및 관리	통계청, 지방자치단체 지방통계사무소
6. 교육·훈련 【 2010. 3~10월 】	○ 교관훈련→총조사 관련 공무원 교육 → 지도원·조사요원 교육 → 자료처 리 관련 교육	통계청, 지방자치단체 지방통계사무소
7. 대국민 홍보 【 2010. 1~11월 】	○ 홍보 세부실시계획 수립→홍보물 제작→홍보실시	통계청, 지방자치단체 지방통계사무소
8. 080 무료전화 운영 【 2010. 10~12월 】	○ 조사 부문별 상담요원 분류 ○ 지도감독을 위한 전화 별도운영	통계청 지방통계사무소
9. 준비 및 본조사 실시 【 2010. 10~11월 】	○ 준비조사 실시(요도·가구명부 작성) →본조사 실시→지도·점검반 편성 및 운용	통계청, 지방자치단체 지방통계사무소
10. 조사표 정리·제출 【 2010. 11월 】	○ 읍면동→시군구→시도 또는 통계 사무소→통계청	통계청, 지방자치단체 지방통계사무소

3) 자료 처리 · 심사단계

업 무 명	주 요 내 용	비 고
1. 자료입력 및 내검 【 2010. 11~12월 】	○ 입력장소 및 장비확보→입력→내검 요원 채용→ 내검실시	통계청
2. 누락·중복 확인조사 실시 【 2010. 12~2011.2월 】	○ 조사계획 수립 ○ 표본 추출 및 현지조사 실시 ○ 결과분석	통계청, 지방통계사무소
3. 자료내검 및 질의조회 【 2011. 1~4월 】	○ 지역단위 내검 및 질의조회 ○ 수준점검	통계청, 지방통계사무소
4. 결과표 집계 및 출력 【 2011. 4월 】	○ Imputation 기법 적용 ○ 결과표 출력	통계청

4) 결과분석 및 공표단계

1. 최종결과 집계 및 공표 【 2011. 5~12월 】	○ 전수집계결과 공표 ○ 표본집계결과 공표	통계청
2. 마이크로 데이터 제공 【 2012. 5월 】	○ 마이크로 데이터 정비 ○ 지방자치단체 제공	통계청 지방자치단체
3. 보고서 발간 【 2011. 9월~ 】	○ 보고서(책자 및 CD-ROM) 발간 ○ 인터넷 DB 구축	통계청

5) 최종평가 단계

1. 평가 및 포상 【 2011. 4월 】	○ 평가 및 포상계획 수립 ○ 평가대회 실시 및 포상	통계청
2. 최종 평가보고서 발간 【 2011. 1~6월 】	○ 평가보고서 발간	통계청

2010년 인구주택총조사 방법론 연구
(요 약)

2007년 11월 23일

통 계 청

2010년 인구주택총조사 방법론 연구 요약

2007년 11월 23일

연구기관 : (사)한국조사연구학회

연구책임자 : 이 건(서울시립대학교 도시사회학과 교수)

공동연구자 : 강영옥(이화여자대학교 사회생활학과 교수)

전명식(고려대학교 통계학과 교수)

이석훈(충남대학교 통계학과 교수)

김규성(서울시립대학교 통계학과 교수)

이명진(고려대학교 사회학과 교수)

서우석(서울시립대학교 도시사회학과 교수)

제1장. 인구주택총조사의 활용성

제1절. 설문조사

이 연구에서는 먼저 인구주택총조사가 어떻게 이용되는지를 알아보기 위해 설문조사를 실시하였다. 조사대상은 공공부문인 중앙부처와 지방자치단체의 공무원과 민간부문인 연구 및 교육기관의 연구자로 하였다. 공공부문에서는 7급 이상 3급 이하 직급의 공무원을 각각의 홈페이지 조직도에서 찾아 표집틀을 만들었다. 여기서 부처별·지자체별로 30명 씩 무작위 추출하여 중앙부처에서 1,240명, 지방자치단체에서 539명을 표본으로 하였다. 민간에서는 관련 학회 및 연구소 명단과 통계청에서 확보한 이용자 명단을 이용하여 90명의 표본을 만들었다. 조사에서는 온라인 조사와 오프라인 조사를 병행하여 총 응답자수는 중앙부처가 177명으로 응답률 14.3%, 지방자치단체가 173명으로 32.1%, 민간인 18명으로 20.0%였다.

조사결과, 지방자치단체가 중앙부처보다는 인구주택총조사를 조금 더 적극적으로 이용하고 있는 것으로 나타났다. 또한, 지방자치단체의 총조사에 대한 요구도 중앙부처에 비해서 더욱 적극적이었다. 요구의 주요 사항은 이용편의, 소지역 통계, 세부 통계, 시의성 있는 통계 등이었다.

이런 측면에서, 먼저 인구주택총조사의 활용성을 높이기 위해서는 지방자치단체에서 필요로 하는 통계에 대해 좀 더 심각하게 고민해 볼 필요가 있다. 기실, 중앙부처는 행정에서 주민등록을 이용하기 때문에 인구주택총조사 통계에 대한 수요가 상대적으로 적은 편이다. 반면, 지방자치단체에서는 지역의 세부사항을 파악하기 위한 통계가 점점 더 필요해지고 있다. 이런 수요에 통계청이 적극적으로 대처하기 위해서는 인구주택총조사에서 지방자치단체의 통계수요를 파악하는 데 좀 더 많은 노력을 기울일 필요가 있다.

한 가지 부연할 것은, 아직 통계에 대해 그다지 많은 경험을 갖지 못한 지방자치단체의 경우 자신들이 어떤 통계를 필요로 하는지 잘 알지 못하는 경우가 있을 수 있다. 이런 경우에 대비하여 지방자치단체의 통계교육을 강

화하고, 통계와 관련된 문제를 함께 상의하며, 그들의 통계문제에 적극적으로 상담해주는 협력체계를 정립하는 것이 바람직하다. 이에 본 연구는 통계청의 통계서비스에 지방자치단체나 중앙부처에서 필요한 통계에 관해 상담할 수 있는 조급은 전문화된 상담창구를 제안한다. 상담창구를 통한 전문화는 통계서비스를 차별화하고, 공공부문 통계서비스의 수준을 질적으로 높이고, 더 나아가 정책에 필요한 통계수요를 파악하는 데 도움을 줄 것이다.

조사와 관련하여 마지막으로, 이 연구는 인구주택총조사 이용자들의 통계요구를 파악하고, 그에 부응하여 새로운 센서스 방법들을 모색하기 위한 것이다. 이런 측면에서, 이 연구는 통계 생산자의 입정보다는 수요자 중심에서 인구주택총조사의 미래를 생각한다는 데 최소한의 의의가 있다. 앞으로 인구주택총조사를 포함한 많은 국가통계는 수요자 입장에서 평가하는 통계의 효용가치가 그 통계의 지속적 생산 가능성을 담보해 주게 될 것이기 때문이다. 따라서 앞으로 통계수요를 파악하는 좀 더 세련된 조사방법을 개발할 필요가 있다는 점도 함께 제시하고자 한다.

제2절. 활용 증진 방안

설문조사결과, 특히 인구주택총조사에 대한 요구 사항들을 바탕으로, 이 연구는 몇 가지 제안을 하고자 한다. 우리나라에서는 행정에서 주민등록자료를 이용하기 때문에 인구주택총조사의 이용확대를 기대하기는 어려울 것으로 보였다. 그렇지만 설문조사는 인구주택총조사에 대한 지방자치단체의 고유한 수요가 있다는 것을 보여주었다. 지역사회의 특성은 주민등록으로 파악할 수 없을뿐더러, 인구주택총조사와 같은 대규모 조사가 아니면 파악하기 힘들기 때문이다. 또한 행정수요의 증가는 행정학이나 사회학 분야에서 지역사회 연구의 증가로 이어지면 이에 따라 학술연구에서도 총조사의 자료에 대한 수요가 늘고 있다. 이러한 소지역 통계의 수요 가운데 특히 총조사의 표본조사에 대한 소지역별 통계에 대한 요구가 가장 많다.

이에 대한 해결 방안으로 이 연구는 공표집계구 설정과 이를 포함한 통계구역체계의 정립을 제안한다. 이를 위해 외국의 통계구역 사례를 살펴보고,

이를 바탕으로 현재 진행되고 있는 우리나라의 통계구역설정을 살펴보았다. 그 가운데 특히 소지역통계 제공에 중심 역할을 하는 집계구를 중점적으로 살펴보았다. 집계구는 통계를 가능한 가장 작은 지역 수준까지 생산함으로써 지방자치단체의 정책에 직접 활용될 수 있는 것이다. 이러한 통계는 단지 행정뿐만 아니라 민간기업의 마케팅, 입지선정, 표본조사의 표집틀 등으로 다양한 활용성을 잠재하고 있다. 또한 국가 중심이었던 학술적 연구가 지역별로 비교연구로 확대될 수 있는 가능성을 제공한다.

한편 집계구 및 기타 통계구역의 기반이 되는 기초단위구는 행정의 또 다른 수요가 있어 주목할 만하다. 도농복합의 성격으로 행정구역을 정의하는데 어려움이 있는 지역들이 많이 나타나고 있다. 이런 지역에서 통계에 바탕을 두고 행정구역을 나눈다면 좀 더 현실적이며 효율적인 행정의 근간을 마련할 수 있다. 즉, 도농복합지역에서 인구주택총조사는 행정구역 지정 및 획정의 근거로 활용될 수 있다.

이러한 작업들은 모두 GIS의 기술적 지원이 또 다른 한 축을 지탱하고 있다. 다시 말해, 전국을 작은 지역 단위로 나누어 그것을 통계의 목적으로 재결합하는 작업은 GIS의 지원이 없다면 구현하기 매우 어려운 것이다. GIS의 이용은 다른 한편, 통계 자료와 지도의 결합을 의미하기도 한다. 이제 지도와 통계 자료는 하나의 DB로 연계·구축되는 것이다. GIS 기술을 기반으로 지역들을 통계적 동질성과 공간적 근접성으로 함께 묶으면, 통계와 지역이 함께 쌍을 이루는 소구역들이 만들어진다. 이러한 소구역들은 지도에 표시하기도 쉽고, 사람들이 이해하기 쉽기 때문에 인구주택총조사의 이용을 촉진시키고 확대하는 데 많은 역할을 할 것이다.

또 하나의 제안은 원자료 이용을 활성화하는 것이다. 통계청의 통계 서비스가 줄곧 발달해왔지만 그 가운데 원자료에 관한 부분이 상대적으로 뒤떨어져 있다. 원자료 서비스는 MDSS와 CD와 같은 매체를 통해 어느 정도 이루어지고 있으며, 비밀보호가 필요한 자료에 대해서는 on-site 서비스도 작은 규모로나마 이루어지고 있다. 그렇지만 아직은 이에 대해 뚜렷한 정책의 방향이 정립 돼있지 않다. 여기서는 자료 형태에 따라 Public Use Microdata와 Confidential data를 구분하여, 서비스를 차별화하는 것을 권한다. 거의 모

든 선진국들이 이와 같은 방식으로 원자료를 서비스하고 있다. 또한 PUM은 무료로, Confidential data에 대해서는 이용 절차와 규정을 까다롭게 적용하며 비용을 책정하는 것이 전반적 경향이다. 이는 통계 이용자가 전문화되어 감에 따라 서비스도 전문화되어 가는 것을 의미하기도 한다.

한편, 원자료 서비스는 정보화 시대의 국가 간 경쟁에서 국민들의 정보역량을 높이는 데 필요한 기초 자원을 국가 제공하는 GOV2.0에 적극적 참여를 의미한다. 이제 통계청은 국가의 가장 기초가 되는 정보를 생산하는 작은 역할에서 벗어나, 이의 유통 및 활용을 장려하여, 국가 경쟁의 초석을 제공하는 큰 역할로 자신의 모습을 바꾸어 나아가야 할 것이다.

제2장. 행정자료 이용

행정자료 이용은 현재 모든 국가가 관심을 기울이는 새로운 센서스 방식이다. 물론 이의 최종 목표는 행정자료로 조사를 대체하는 것이다. 그렇지만 북구유럽을 제외하고는 아직 행정자료 기반 센서스를 본격적으로 실시하는 나라는 거의 없다. 북구유럽에서도 덴마크와 핀란드는 이미 오래 전부터 행정자료로 모든 센서스 통계를 대체했지만, 스웨덴은 지속적으로 행정자료로 센서스를 대체해 나아가고 있으며, 노르웨이는 2011년에 완전 대체를 목표 현재 거주주소(dwelling household project)를 정비하고 있다.

2000년대에 들어 우리나라에서도 행정자료로 센서스를 대체하자는 의견이 빠르게 전파되고 있다. 여기에는 1962년 제도화된 우리의 주민등록제도가 이제 상당히 안정되어 행정의 전 분야에서 사용되는 것이 가장 큰 이유다. 또한 1990년대 말과 2000년대 초에 부동산 관련 행정자료, 특히 건축물대장 자료의 정비도 행정자료로 센서스를 대체할 수 있다는 의견에 큰 몫을 차지한다. 이런 의견들이 얼마나 현실적으로 가능하지를 알아보기 위해 인구주택총조사자료와 주민등록자료를 전국, 시도, 시군구, 읍면동 수준에서 비교해 보고, 2개의 지역에서 2004년에 실시한 인구주택총조사 시범조사 자료와 같은 시기의 주민등록자료를 개인 수준에서 비교해 보았다. 또한 건축물과 관련

있는 행정자료들과 이들 사이의 연계 가능성에 관한 기존 연구들을 발굴하여 정리했다.

그 결과 지금 당장은 행정자료로 인구주택총조사의 조사 항목에 대한 통계를 대체할 수 없다는 사실을 알 수 있었다. 일단 인구주택총조사와 주민등록자료의 인구 및 가구-세대 차이가 컸으며, 행정구역별 집계수준이 작아질수록 총조사와 등록 자료의 차이는 커지고 편차도 커졌다. 인구의 경우 전국 수준에서는 전체적으로 4% 정도의 차이가 났지만, 읍면동 수준에서는 전체 3,572개 동 가운데 1,239개(34.7%)가 10% 이상 차이가 났으며, 30% 이상 차이가 나는 읍면동이 40개 가량 있었다. 또한 이상값을 제외하더라도 가장 큰 차이는 60% 이상이었다. 가구-세대의 차이는 이보다 더 커서 전국 수준에서는 10% 정도의 차이가 났고, 읍면동 수준에서는 90% 이상 차이가 나는 지역도 있었다. 이와 같은 현상은 연령별 차이에서도 마찬가지였다. 연령별 분석에서 가장 두드러진 결과는 15-24세 사이의 연령층에서 주민등록과 총조사 자료 사이의 차이가 큰 것이다. 이는 젊은 시절에 대학이나 직장으로 거주지를 옮기는 경우가 많은데 이때 주민등록의 주소는 바꾸지 않기 때문에 나타난 것으로 이해할 수 있다.

이 연구에서 나타난 주민등록과 인구주택총조사의 차이는 앞으로 주민등록으로 총조사 인구를 대체하기 위해서는 이러한 차이를 보정할 수 있는 모형의 개발이 필요하다는 것을 제시한다. 이를 위해서는 먼저 인구에서는 인구주택총조사와 같은 시점의 주민등록 자료를 사용하고, 총조사와 같은 집계 방법으로 집계하여 비교해 보아야 할 것이다. 더 나아가 이러한 자료를 시계열로 비교해 보아야 할 것이다. 다음은 이러한 비교 결과를 토대로 보정 모형을 개발해야 한다. 이때 연령별 특성과 지역의 특성은 모형에 중요한 역할을 할 것으로 보인다.

보정 모형 개발의 최종 단계에서는 개인 수준에서 두 자료를 비교하는 과정을 거쳐야 한다. 이러한 작업을 통해 개인 수준에서 발생하는 주민등록과 실제 거주 차이에 대한 좀 더 확실한 이해가 있어야 신뢰할 수 있는 보정 모형을 만들 수 있을 것이다. 더욱이 이 연구에서 다룬 자료가 읍면동 수준이 가장 작은 집계 수준이어서 집계구와 같이 이보다 작은 지역의 집계를

내기 위해서는 좀 더 세밀한 모형이 필요하기 때문이기도 하다. 이를 위해서는 2008년 시험조사서부터 본 조사, 더 나아가서는 본 조사 이후에도 여러 차례에 걸쳐 주요 지역들, 특히 지속적으로 두 자료의 차이가 많이 나는 지역부터 표본을 추출하여 개인 수준에서 직접 주민등록과 조사 자료를 비교하여야 할 필요가 있다. 이때 적어도 가구주의 이름, 가구원의 생년월일, 거처의 주소를 함께 조사하여 주민등록과 직접 비교가 가능하도록 해야 할 것이다. 또한 가능하면 가구원의 이름을 두 자리만이라도 받을 수 있으면 더욱 좋을 것이다.

건축물 관련 행정자료는 인구주택총조사에서 가구와 주택에 관한 사항을 대체하는 데 이용될 수 있다. 그렇지만 현재는 건축물 관련 행정자료들 사이의 문제, 자료 자체의 정확성, 서로 다른 분류 체계 등으로 사실상 이용이 불가능하다. 가장 시급히 개선할 문제는 현재 사용하는 주소인 필지번호(지번)다. 현재 필지 중심의 주소 체계로는 건물에 고유한 번호를 부여하기 어렵다는 점이다. 이는 단지 통계의 문제뿐만 아니라 이미 도로 중심으로 건물에 번호를 부여하는 새주소가 만들어지고 있다. 또한 새주소법이 국회를 통과하면 공식적으로 행정에서도 새주소를 사용하게 되어 건물의 고유한 번호를 사용하기는 점점 나아질 것이다.

그렇지만 새주소로도 행정자료에서 총조사의 가구정보를 취득하는 데는 한계가 있다. 건물 내부의 거처별로 고유한 번호가 있어야 이를 바탕으로 같은 거처에 사는 사람들을 인식할 수 있고, 그래야만 이들을 하나의 가구로 정의할 수 있기 때문이다. 스웨덴과 노르웨이가 행정자료 기반 센서스로 나아가는 데 어려움을 겪고 있는 데는 바로 거처별로 고유한 주소를 갖지 못한 것인 큰 원인이었다. 이런 이유에서 노르웨이는 거처별로 주소를 부여하는 거처주소프로젝트를 실시하여 진행 중에 있으며, 이를 바탕으로 2011년에는 완전히 행정자료 기반의 센서스가 될 것을 기대하고 있다.

새주소가 정착되어 건축물 관련 행정자료를 서로 연계할 수 있다고 하더라도, 만약 개별 자료가 부정확하다든지 누락되어 있다면 이 또한 행정자료로서 사용가치가 매우 낮아진다. 또한 건축물 관련 항목의 분류체계나 항목 자체의 의미가 다르면 이 또한 사용하기가 어렵다. 따라서 건축물 관련 행정자

료로 총조사를 대체하기 위해서는 개별 행정기관에서 정확히 자료를 입력하고, 적절한 주기로 갱신해야 하며, 더 나아가서는 같은 분류체계와 같은 용어를 사용해야 한다. 그렇지만 이 모두가 통계청이 할 수 있는 일이 아니기 때문에 모든 부처가 공히 사용할 표준안을 만들고 이를 사용하기까지는 그다지 쉽지 않을 것이다.

이런 어려움에도 불구하고, 가구와 주택부분의 인구주택총조사 항목 가운데 건축물 관련 행정자료와 통계청의 누적된 인구주택총조사자료를 함께 이용하면 인구주택총조사의 응답부담을 줄일 수 있는 부분이 있다. 지난 2005년 총조사에서 가구와 주택에 관한 항목은 통계청의 DB에 들어가 있다. 2005년 이후 신축하거나 개축한 경우와 변경하거나 부순 경우 이외에는 이들은 변하지 않았다. 이러한 항목은 가구에 관한 사항 가운데 사용 방수, 주거 시설 형태, 거주층, 난방시설이며, 주택에 관한 사항은 거처의 종류 및 건물층수, 연건평, 대지면적, 총방수, 건축년도, 편익시설수가 있다. 먼저 건축물대장과 건축허가신청서 등을 이용하여 건축물 신축, 개축, 변경, 붕괴 등을 점검한다. 변동이 없는 건축물에 대해서는 과거의 자료를 그대로 사용하고, 변동 사항이 있을 경우도 조사에서는 변동된 사항을 확인하여 수정하면 응답자 부담을 줄일 수 있을뿐만 아니라, 조사비용도 줄일 수 있을 것이다. 물론 이러한 과정은 시험조사를 통해 실제로 어떤 문제가 있는지를 파악하고, 특히 조사원 교육에서 발생할 문제를 미리 점검하는 것이 필요할 것이다.

행정자료 이용에서 최종의 목표는 행정자료로 모든 센서스 통계를 생산하는 것이다. 이를 위해서는 건축물 관련 행정자료에서도 논의하였듯이 많은 작업이 필요하며, 더욱이 그러한 작업은 통계청이 혼자서 할 수 있는 것이 아니다. 따라서 행정자료 기반의 센서스를 추진하는 자세는 길게 내다보고, 관련 행정부처에 협조를 구하고, 더 나아가 국가적 차원에서 협조체계를 구축하고, 행정절차에 통계가 스며들도록 하는 것이다.

앞으로 인구주택총조사에서 행정자료를 활용하기 위해서, 더 나약 행정자료 기반의 센서스로 옮겨 갈 수 있기 위해 지금 해야 할 사안들은 다음과 같다.

- 통계청이 확보한 기존 자료를 최대한 이용할 수 있는 체계구축
- 보완적으로 활용 가능한 행정자료의 발굴 및 이용체계 구축
- 새주소체계 활용 및 거처별 번호 부여 작업
- 수치지도와 거처별 번호 연계
- 조사자료에 거처별 번호 입력
- 조사자료와 행정자료와 비교 연구

제3장. 순환(총)조사

순환(총)조사에 대해서는 프랑스의 순환총조사와 미국의 순환조사를 살펴 보았다. 아직 순환(총)조사가 시행 중에 있으며, 총체적인 결과가 나오지 않아 순환(총)조사방법에 대해 평가하기는 이르다. 현재까지 두 국가에서 시행해 온 것을 토대로 우리나라의 인구주택총조사에 어떤 시사점을 주는지 살펴 보았다.

먼저 순환조사를 프랑스와 같이 전수조사에 적용할 것인지 아니면 미국과 같이 표본조사에 적용할 것인지를 검토해야 한다. 결론부터 이야기 하자면, 우리 실정에 전수조사보다는 표본조사가 더 적합하다. 왜냐하면, 현재 인구주택총조사를 시행하는 데 문제가 있는 것은 아니기 때문에 5년에 한번 하는 전수조사를 굳이 바꿀 필요는 없을 것이다. 다만 5년에 한번 결과가 발표되어 통계의 시의성이 떨어지는 문제는 고려해 볼 만하다. 특히 지방자치단체가 총조사 통계에 대해 더 적극적이며, 그 가운데도 표본조사로 만들어지는 소지역 통계에 대한 수요가 크다. 이런 측면에서 표본조사를 순환조사로 하면 매년 소지역 통계를 생산할 수 있을 것이다. 이때 소지역의 수준이 문제인데, 매년 생산하는 통계는 시군구 수준, 5년을 합한 통계는 읍면동 수준 까지 생산이 가능할 것이다.

지금 10% 표본조사를 그대로 시행한다고 가정하고 먼저 읍면동 수준에서 추출한다면, 전국의 읍면동이 3,584개니 표본으로 뽑히는 읍면동은 약 360개가 된다. 5년 동안 매달 조사를 한다면, 60개월의 조사가 되고, 한 달에 6개의 읍면동을 조사하게 된다. 가구로 보면 10% 표본이 150만 가구니, 한 달에

는 25,000 가구를 조사하는 셈이다. 이를 종합하면, 평균적으로 한 달에 한 동에서 420 가구씩, 6개 읍면동을 조사하면 현재와 같은 10% 표본을 조사할 수 있다.

그 다음 문제는 시군구 수준에서 대표성 확보를 몇 년 주기로 하느냐는 것이다. 복잡한 경우는 인구수와 행정구역 수가 많은 경우다. 예를 들면 서울에는 25개 구가, 경기도에는 31개 시군구가 있다. 이러한 시도에서 시군구 전체를 대표하는 자료를 매년 생산할 것인지, 아니면 매년 일부 시군구만 대표성 있는 자료를 생산하고 5년을 합쳐 시군구를 대표할 자료를 생산할지, 아니면 2-3년의 주기로 할지 등을 정해야 한다. 현재로는 이에 대해 어느 것이 좋다고 정할 수는 없으나, 표본조사를 순환조사로 할 경우에는 먼저 결정해야 할 사안이다.

다음은 방법론의 개발이다. 특히 표본틀(sampling frame)이 계속 변하기 때문에 이에 대한 대책이 있어야 한다. 이에 대해서는 현재와 같이 주민등록의 인구변동을 가져다 인구를 추정하는 수밖에 없다. 물론 이를 위해서는 행정부와 지방자치단체에서 정기적으로 주민등록의 변동 사항을 통계청에 통보해 주어야 한다. 따라서 이 부분에서도 행정자료의 이용이 필수적인 사안이 될 것이다.

그렇지만, 현재 인구이동통계를 내는 경우와는 고려해야 할 사항이 다르다. 인구이동통계는 발생한 현상, 즉 과거의 것을 통계에 반영하는 것이다. 반면, 표본설계는 미래의 것을 준비하는 작업이다. 적어도 조사시점의 12달 전에는 1년 조사의 표본설계가 완료되어있어야 하며, 이상적인 경우는 60개월 전체에 대한 표본설계가 있어야 할 것이다. 이렇게 미래의 지역별 인구를 추정하기 위해서는 지역에 대한 예측 가능한 사건, 예를 들면 재개발인 택지개발 등에 대한 고려를 해야 할 것이다. 또한 예측이 불가능 사건, 예를 들면 재해나 천재지변 등이 발생하여 인구가 감소(증가)하는 경우 어떻게 대응할 것인지에 대한 절차 및 방법에 대한 사전의 준비가 있어야 할 것이다.

이외에도 심각한 방법론의 개발이 요구되는 사안들이 많다. 특히 현재 표본조사에서 사용하는 표집단위(sampling unit)인 조사구로 대표성을 확보할 수 있는지는 의문이다. 조사구들 사이의 표본 수 차이가 너무 크기 때문이

다. 따라서 표본조사를 순환조사로 할 경우 현재의 조사구보다는 좀 더 일관된 크기의 새로운 표집단위의 개발해야 한다. 이는 지금까지 통계청이 해온 모든 표본조사의 근간을 바꾸는 작업이다. 경제활동인구조사나 가계조사와 같은 통계청 조사뿐만 아니라 통계청으로부터 표본의 협조를 받는 다른 부처의 표본조사에도 큰 변화가 따른다는 점도 고려해야 할 것이다.

이와 더불어 모수추정에 대한 새로운 방법론 개발이 필요하다. 연차 자료의 누적방법에 대해 이동 평균법과 가중 방법을 개발해야 하며, 5년 누적 데이터를 이용하여 인구수 및 가구수를 추정하는 방법을 개발하여야 한다. 또한 가구와 인구가 이동하는 상황에서 모수를 추정하는 모형을 개발해야 한다. 또한 공표하고자 하는 결과의 지리적 및 인구통계적 범위를 결정해야 한다.

다음으로 고려할 문제는 상시조사체제가 되기 때문에 이 조사에만 전념하는 전문요원이 필요하다는 것이다. 예를 들면 하루에 10 가구씩 방문조사를 한다고 하면 한 달에 한 조사원이 20일을 조사에 투입한다고 보았을 때 200 가구를 방문하여 조사할 수 있다. 이럴 경우 125명의 표본조사 전담 조사원이 필요하다. 이것은 인력에 대한 예산과 함께 지금 한창 문제가 되는 비정규직 고용과도 직접 관계가 있다. 이런 점들을 감안해보면, 전체적인 예산은 현재의 표본조사에 비해 많이 들 것이며, 특히 초기에 방법론 개발과 조사체계를 안정되게 구축하는 데 비용이 많이 들 것이다. 더불어 여러 차례의 시험조사와 본조사를 비교연구하고, 이러한 과정을 여러 차례 거쳐 적절한 방법론을 개발하는 장기간에 걸친 사업으로 진행하여야 유용한 통계가 생산될 수 있다는 점도 염두에 두어야 할 것이다.

간략하게 2015년 인구주택총조사의 표본조사를 순환조사로 대체한다고 할 때 이에 대비해 해야 할 최소의 사항들을 열거하면 다음과 같다.

- 2010년 표본조사와 비교를 위한 2009년, 2010년, 2011년의 조사
- 2009년 조사를 위한 여러 차례의 2008년 시험조사 및 평가
- 조사구보다 작으며, 어느 정도 일정한 규모의 표집단위 설정 및 구축

제4장. 제언

이 연구의 주된 관심은 우리나라의 센서스가 앞으로 어느 방향으로 가야 할지를 탐색하는 것이다. 현재 우리가 직면한 상황은 행정자료 기반의 센서스와 순환조사 가운데 어느 쪽을 선택할 것인지를 결정하는 것이다. 물론 이 연구 하나로 앞으로의 방향을 결정하는 것은 매우 위험하다. 따라서 이 연구에서는 가능하면 다각적이며 정확한 근거를 가진 제안을 하여 의사결정자들이 의사결정에 도움을 주도록 하는 것이 최선일 것이다.

먼저 행정자료 기반의 센서스는 많은 국가들이 현재의 센서스를 대체할 미래의 센서스 방법으로 생각하고 있다. 현재는 북구유럽의 몇 국가들이 실시하는 정도에 불과하지만 다른 나라들도 행정자료 기반의 센서스를 향해 여러 가지 방법들을 시도해 보고 있다. 그 가운데 행정자료를 조사에서 보완적으로 사용하는 경우가 가장 일반적이다.

한편 순환조사는 단지 두 나라가 시행하고 있을 뿐이어서 행정자료 기반의 센서스에 비해 그다지 보편적인 것은 아니다. 행정자료 기반의 센서스는 이 방법을 20년 이상 시행한 국가들이 있기 때문에 어느 정도 검증도 되었고, 방법론을 배울 수도 있다. 이에 반해 순환조사는 아직 최종의 결과가 나오지 않은 상태여서 조사에 대한 검증이 이루어지지 않았고 독자적으로 개발해야 할 방법론적 과제가 많다.

이런 국제적 경향을 보면, 우리가 왜 이 시점에서 센서스의 향방을 생각해 하든지부터 살펴볼 필요가 있다. 먼저 우리나라에서 센서스 시행에 아직 커다란 문제가 생긴 것은 아니다. 독일이나 네델란드처럼 응답거부가 많은 것도 아니고, 호응도 그다지 나쁘지 않다. 특히 지난 총조사에서 홍보는 조금은 축제 분위기를 만들었던 것 같다. 이런 측면에서 일반 국민들이 생각하는 센서스는 현재로는 그다지 큰 문제는 아닌 것으로 보인다.

반면 총조사의 비용증가와 결과의 유용성에 대한 비판은 조금은 심각하게 다가오고 있는 것으로 보인다. 비용증가는 인구증가로 인한 조사비용 증가, 노인가구와 단독가구의 증가로 인한 조사비용 증가, 지방자치단체의 통계기능 약화에 따른 비용 증가 등 여러 원인이 있지만 그러한 원인은 어쩔 수

없는 것이고 그 때문에 인구주택총조사를 비판하는 것은 아니다. 여기에는 ‘인구주택총조사에 그 많은 비용을 들여서 어디에 쓰는지’로 대표될 수 있는 비판, 즉 비용증가에 따른 유용성에 대한 회의가 있기 때문이다. 이런 측면에서, 우리가 인구주택총조사의 앞날을 생각할 때는 비용절감의 측면과 유용성 확대의 두 면을 함께 고려해야 할 것이다.

행정자료 기반의 센서스를 생각해 보면, 이 방법이 완전히 정착될 경우 장점이 단점에 비해 훨씬 우월하다. 다만, 우리나라가 과연 이러한 통계생산의 체계를 갖출 수 있을지에 대해 확신이 서질 않는다. 왜냐하면, 이러한 체계를 갖추고 정착시키는 데는 통계정보보다 훨씬 넓은 범위, 때로는 많은 행정부처와 지방자치단체, 더 나아가 국가의 모든 기관과 국민들 적극적인 협조내지는 동의가 필요할 수도 있기 때문이다. 더욱이 이런 체제가 제대로 정착되는 데 얼마나 오랜 시간이 걸릴지 아무도 예측할 수 없는 것도 문제다.

이에 반해 순환조사는 통계청이 혼자서 수행하기에 상대적으로 수월하다. 이런 측면에서 당장 순환조사에 대한 욕심이 생기는 것도 어쩔 수 없는 인지상정이다. 특히 요즈음처럼 인구주택총조사에 대해 시의성 있는 통계, 소지역 통계의 요구가 커진 상황에서 순환조사를 시행하면 이러한 요구에 부응할 수 있을 것도 같으며, 더욱이 통계청이 독자적으로 통계 수요에 적극적으로 대처하는 모습을 보여 줄 수 있어 매력적이다.

이렇게 보면, 행정자료 기반의 센서스인가 아니면 순환조사인가는 통계수요에 대한 장기적 대응인가 단기적 대응인가를 선택하는 문제로 보인다. 일단 이렇게 장기와 단기로 나누면, 그 다음은 이 둘을 단계적으로 모두 시행해 볼 수 있을 것 같기도 하다. 즉, 단기적으로 순환조사를 시행하고, 장기적으로는 행정자료 기반의 센서스로 나아가는 방안이다. 그렇지만 이 방안은 이중 비용 부담에다, 중국에는 순환조사에서 개발한 방법을 폐기하게 되는 문제가 있다.

이 시점에서 순환조사를 표본조사에만 국한시켜 표본조사는 순환조사로 지속하되, 행정자료 기반의 센서스를 향해 전반적인 통계체계를 바꾸는 방안을 생각해 보자. 지금까지 표본조사는 전수조사와 함께 시행되면서 마치 둘은 뗄 수 없는 관계에 있는 것처럼 되어 왔다. 이 둘의 관계를 보면 실제 조사

내역은 크게 관계가 없지만, 차후에 다른 표본을 설계하는 데 이들은 함께 사용된다. 예를 들면 경제활동인구조사의 표본설계에 표본조사 자료가 사용되며, 이때 전수조사 자료는 모집단의 역할을 하고, 가중치를 부여하는 데 중요한 역할을 한다. 물론 완전한 행정자료 기반 센서스에서는 표본조사가 불필요하다. 단지 그 완전함을 확보하는 데 많은 시간이 들며, 그 사이에도 통계는 생산되어야 하기 때문에 여전히 표본조사가 필요하다. 이런 측면에서 당분간은 표본조사가 있어야 할 것으로 보인다.

적어도 당분간 표본조사가 있어야 한다는 것을 기정사실로 받아들이면, 이를 순환조사로 바꿀 것인가를 고려해야 한다. 앞서 이야기 방법론의 문제나 비용 등의 문제를 제외하더라도 또 다른 문제를 생각해 볼 수 있다. 표본조사가 60개월짜리 순환조사가 되면 현재 진행되는 다른 조사들과 어떻게 조화를 이룰지의 문제다. 경제활동인구조사와 비슷한 내역을 조사하는 것은 특히 문제가 될 수 있다. 이런 측면에서 다른 조사들에 대한 조정 작업이 있어야 할 것이다. 더욱이, 이런 모든 문제들을 고려하면, 표본조사의 시의성과 유용성에 관한 문제는 다른 방식의 소지역 통계로 대체가 가능한지도 검토해 보아야 할 사항이다.

이제 우리는 한국의 통계생산 전반을 놓고 고려하게 되었다. 다시 원점으로 돌아가야 할 때다. 다시 한번 현재의 상황을 그려보면, 행정에서 주민등록자료의 정비와 활용은 행정자료 기반의 센서스에 대한 요구로 이어지고 있다. 이를 거부할 명분이나 대안은 없다. 따라서 우리는 행정자료의 이용을 확대하며 행정자료 기반 센서스로 나아가야 한다. 행정에서 주민등록이 차지하는 역할의 비중과 활용 가능성이 소지역 통계에 대한 필요성보다는 훨씬 더 범위가 넓고 파급효과도 크기 때문이다.

이를 전제로 이제 통계청은 행정자료를 이용하기 위해 개별 행정부처에 자료를 요구하는 수준에서 벗어나 더 공격적으로 행정자료를 통계적 목적으로 활용할 수 있도록 행정업무의 설계와 집행에 한 부분을 담당할 필요가 있다. 어차피 개별 행정부처에서도 행정을 통한 통계집계가 점점 더 필요해지기 때문이다. 이를 발굴하고 알려주는 것도 통계청의 역할이다. 본 연구진은 통계청이 앞으로 이러한 역할을 하지 않으면 다른 부처가 이를 대신할

것이며 결국 그들이 필요한 통계기능을 가져가고 통계청의 역할은 줄어들 것이다. 예를 들면, 행정자치부와 정보통신부는 행정정보공유센터를 만들어 자신들의 정보를 연계하는 작업을 시작했다. 또 건설교통부와 행정자치부는 부동산정보관리센터를 통해 종합부동산세를 책정하기 위해 각종 대장을 연계하기 시작했다.

이런 측면에서 보면, 행정자료 기반의 센서스는 선택이 아니라 반드시 가야할 수밖에 없는 길이다. 표본조사를 순환조사로 할 것인가 아닌가는 그에 비하면 훨씬 작은 규모의 논의다. 따라서 수요가 있고 충분히 비용을 지불할 수단 있다면 시행할 수도 있는 것으로, 의사결정자들의 판단에 달린 문제다. 이에 비해 행정자료 기반의 센서스는 의사결정의 문제가 아니다. 이것은 시대의 주어진 상황이며, 통계청은 그에 비추어 미래에 설 자신들의 역할을 정립하는 비전을 갖는 것이 중요하다.

통계의 필요성에 비추어, 행정자료 기반의 자료 연계는 통계청이 주도하는 것이 가장 바람직하다. 그렇지 않으면 각종 행정정보가 개인수준에서 연결되어 개인을 통제하는 수단이 될 수 있기 때문이다. 이럴 경우 국민들이 행정정보의 연계 자체를 거부할 수도 있어 여러 행정자료들을 연계하는 시기가 매우 늦춰질 수 있다. 통계청은 개인수준에서 연계된 개인정보 자체에는 관심이 없으며, 이들을 집계하여 국가와 사회에 필요한 통계를 생산하는데 관심이 있다. 따라서 개인정보를 관리하는 하위 이를 개인수준의 정보로 활용하지 말아야 하는 서로 어긋날 것 같은 역할에는 통계청이 가장 적합하다. 각종 개인정보들을 개인수준에서 연계하여 종합적인 개인정보의 집합체를 구축하지만 이를 반드시 집계하여 통계로만 외부에 제공하는 것이 통계청 본연의 임무이기 때문이다. 이런 차원에서, 통계청은 개인정보의 연계와 이의 보호를 주도함으로써 개인정보보호에 민감하면서 동시에 그러한 자료를 모아 행정에 필요한 통계를 생산하고, 국민에게 국가통계를 제공하는 중심 역할을 선도해 가야 할 것이다.