

인구센서스와 행정자료의 활용

김형석 · 양경진 · 강영민* **

1. 서론

오늘날 인구센서스는 수요와 공급의 환경 변화로 조사 패러다임의 변화가 요구되고 있다. 사회가 복잡해지면서 보다 다양하고 상세한 통계 수요는 늘어나는 반면, 조사하기 힘든 유형의 가구(1인 가구, 노인가구 등)나 조사 비협조 가구는 크게 증가하여 정확도가 갈수록 떨어지고 있다¹⁾. 여기에다 조사비용은 기하급수적으로 늘고 있고²⁾, 16만 명의 조사원 및 지도원을 관리하는 읍·면·동의 통계 조직은 정부의 기능전환³⁾ 시책으로 대부분 없어져 인구센서스의 조사체계가 크게 취약해졌다.

인구 규모를 헤아리는 일은 예로부터 국가경영의 중요한 수단중의 하나였다. 우리나라에서 부녀자, 노인, 아동, 천민 등과 같은 특정집단의 인구를 빠뜨리지 않고 모든 인구를 통계 목적으로 조사한 최초의 인구조사는 1925년에 실시한 제1회 인구센서스이다. 센서스에서 인구규모 변동의 가장 주된 요인인 출산력에 관심을 갖고 조사하기 시작한 것은 개발기인 1966년 센서스부터이며, 1990년 센서스까지 지속되었다. 하지만 1980년대 후반 우리나라가 인구변천을 마무리(post-transition; 저출생·저사망으로 인구가 안정적으로 성장)함에 따라, 인구규모를 파악하는 인구센서스의 전통적인 역할에 변화가 나타나기 시작했다. 소수 의견이기는 하지만 막대한 예산과 인력을 들여 센서스를 실시하기보다는 그동안 꾸준히 질적으로 개선되어 온 주민등록자료의 사용을 주장하는 사람도 생겨났다.

먼저 우리는 인구센서스의 역할 변화에 주목해야 한다. 1990년대부터 인구문제는 양적

* 통계청 인구조사과

** 이 글의 내용은 저자 개인 의견임을 알려드립니다.

- 1) 인구센서스 사후조사(post-enumeration survey) 결과를 보면, 순누락률이 1995년 1.3%에서 2000년에는 1.6%로 확대되었다.
- 2) 센서스 비용(당해연도 및 앞뒤 연도 관련예산 모두 포함)이 1980년 43억원에서 1990년 244억원, 2000년에는 969억원으로 증가하였다.
- 3) 정부의 기능전환은 1998년 2월 국민의 정부 100대 과제로 선정되었으며, 이후 행정자치부에서는 1999. 2월부터 “읍·면·동사무소 기능전환 기본계획”에 의거 읍·면·동 사무소를 폐지하여 주민자치센터로 전환을 추진해 왔다. 2003년 현재 전체 3,516개 읍·면·동중 2,710개 읍·면·동의 통계기능이 폐지되었는데, 이 중 경기도(491개)와 광주광역시(85개)는 전체 읍·면·동에서 통계기능이 폐지되었다.

인 측면에서 질적인 측면으로 중심이 이동해 왔다. 아동, 여성, 노인 복지의 문제를 비롯해 가족, 주거 등 삶의 질에 관련된 통계 수요가 크게 늘었으며, 읍·면·동 단위 이하의 소지역 통계에 대한 수요도 지리정보시스템의 발전으로 더욱 늘어나고 있다. 오히려 선진 복지사회 진입과 정보 기술의 발전 수준에서 볼 때 인구센서스는 전 국민에 대한 개별정보를 가지고 있다는 점에서 21세기 정보의 보고(寶庫)로 새롭게 인식되어야 한다.

세계의 인구센서스 실시국가(2000년 라운드 센서스의 경우 세계 229개 국가 및 지역 중 213개에서 실시)들은 공히 센서스에 대한 국민의 비협조와 비용증가 문제로 고민하고 있다. 이러한 센서스의 당면과제와 센서스 자료의 개별성(individual enumeration)에서 우리는 센서스 패러다임의 변화 가능성에 주목해야 한다. 그런데 이러한 변화 가능성은 최근 정보통신 기술의 발전으로 자료의 축척, 접근 및 활용도가 크게 개선된 행정자료⁴⁾에서 찾을 수 있을 것이다.

근자에 행정자료는 등록 및 신고 수준의 개선으로 내용이 크게 개선되고 있을 뿐만 아니라, 데이터베이스화되어 정보 연계 시스템으로 구축되고 있어, 통계에 보다 폭 넓게 활용할 수 있는 여건이 빠르게 조성되고 있다. 예를 들어, 주민등록, 건축물, 국세자료, 국가공인 자격증(701종) 등이 전산화되었으며, 4대 사회보험(국민연금, 건강보험, 고용보험, 산재보험)과 5대 민원(주민, 부동산, 자동차, 기업, 세금) 데이터베이스 공동이용 시스템(G4C)이 이미 구축되었다.

물론 행정자료는 포괄범위, 기준, 목적 등 자료의 작성과정이 조사통계와는 근본적으로 다르다. 하지만 두 가지 자료원은 자료의 개별성이라는 점에서 공통점을 갖는다. 더구나 행정자료가 개인정보 형태로 전산화되어 네트워크로 구축되어 가고 있는 반면 전통적인 조사통계는 국민의 조사 비협조와 조사비용의 증가라는 문제를 안고 있어, 조사통계에서 행정자료의 활용을 본격적으로 모색하지 않으면 안 될 상황이 되었다. 현재 행정자료는 일부 직접 보고통계로 작성되거나 분석통계로 재차 가공되어지기도 하지만, 대부분 조사통계의 작성과정에서 표본틀이나 내용 검토시 참고자료 등으로 활용되는 정도이다(<부록 1> 참조). 통계청에서는 그동안 사업체 통계를 중심으로 행정자료 활용을 확대하기 위해 노력해 왔다. 2000. 4월에는 도소매업 및 서비스업 통계조사의 모집단으로 국세청의 사업자등록 신고자료 및 부가가치세 신고자료를 활용할 수 있도록 추진해 왔으나 국세청의 비협조로 결실을 맺지 못했다.

만약 인구센서스와 행정자료가 성공적으로 접목된다면 경제적으로나 통계적으로도 커다란 변화가 예상된다. 우선 행정자료의 전부 또는 일부를 대체 사용함으로써 국민의 응답 부담을 경감하고 조사비용도 절감할 수 있다. 행정자료의 활용으로 조사누락 및 오류를 보완함으로써 인구센서스의 정확성 및 효율성도 제고시킬 수 있다. 또한 양 자료의 접목

4) 여기서 행정자료는 기업체, 학교, 병원, 협회, 공단, 지자체 및 정부에서 행정업무에 수반되어 작성되는 신고, 보고, 등록 자료를 의미한다.

으로 다양한 고급 통계정보의 생산이 가능하다. 즉, 인구센서스와 행정자료를 개별수준에서 접목하면 그동안 흩어져 있던 수많은 정보들을 개인, 가구, 사업체별로 연계하여 종합 통계정보로 사용할 수 있게 된다. 또한 수시로 자료가 신속하게 작성되어 인구센서스를 매년 실시하는 효과를 갖게 된다. 아울러 이렇게 연계된 자료가 시계열로 축적되면, 종단면 분석(longitudinal analysis)이 가능해 인구센서스의 패러다임에서 변혁을 가져오게 된다. 물론 행정자료를 활용하는 데에는 단점도 많이 있다. 행정자료에 많이 의존하면 자료 수집에 융통성이 없어지고 인구의 경제활동 특성 같은 일부 자료는 정확도에서 문제가 생긴다. 특히 개인정보의 보호라는 측면에서 볼 때, 이러한 자료통합이 사회적 반감을 불러 일으킬 수도 있다.

이 글에서는 인구센서스의 행정자료 활용 현황에 대해 살펴보고, 이를 토대로 인구센서스와 행정자료와의 통합 모형을 제시한 후, 이 모형 안에서 다시 주민등록자료와의 연계 가능성에 대해 살펴보기로 한다.

II. 인구센서스의 행정자료 활용

1. 한국의 2000년 인구센서스와 행정자료

인구센서스에서 행정자료 활용은 지도작업에서부터 조사구 설정, 조사개념, 현장조사, 자료분석에 이르기까지 직·간접적으로 다양하게 이루어진다. 조사개념이나 현장조사에 가장 폭넓게 행정자료를 활용하는 분야는 주택항목이다. 우선 주택항목들은 주택의 유형(단독주택, 아파트, 다세대주택 등)에서부터 건축년도, 취사연료, 난방시설 등에서 행정적인 개념을 도입하고 있다. 따라서 관련 행정자료(예를 들면, 지역별 주택수, 중앙난방 호수, 건축년도별 주택수 등)들은 조사과정이나 결과분석시 참고자료로 사용된다.

조사구 설정에서는 지역별 호텔이나 사회복지시설수, 주민등록 세대수 같은 행정자료가 중요하게 취급된다. 조사항목에서도 학교급별 학생수, 아동보육시설수, 학원수 같은 행정 통계는 조사항목 설계에서부터 자료분석에 이르기까지 참고자료로 널리 사용된다.

인구센서스에서 행정자료 활용은 점점 더 영역이 확대되고 있다. 지난 2000년 센서스에서는 “공동주택기본정보 사전수집 및 활용”이라는 독특한 조사방법을 개발하였다. 이는 아파트와 같이 주택사정이 유사한 공동주택의 경우 관리사무소나 그 주택사정을 가장 잘 아는 주민을 대상으로 아파트에 대한 공통 주택정보(연건평, 건축년도, 주거시설, 편익시설, 총방수, 상수도 시설 등)를 준비조사 기간 중에 수집하여 조사표에 연필로 기입하였다가

현장조사시 활용하는 조사방법이다. 이 방식의 도입으로 조사원의 업무량 및 가구의 응답 부담이 줄어 든 것은 물론 동 항목에 대한 조사의 정확도도 크게 높아졌다.

한편 인구센서스에서 무엇보다 주목해야 할 행정자료는 주민등록자료이다. 주민등록자료는 1962년에 제정된 주민등록법에 근거하여 시·도에서 신고통계로 작성하고 있다. 주민등록부(전입신고)는 인구의 경제·사회적인 특성항목과 가구에 관한 사항을 많이 포괄하고 있으나, 최초 기재이후 거의 갱신하지 않는다는 문제가 있다. 또한 부동산 매매, 자녀의 학군배정, 부양가족 혜택, 취업, 투표권 등의 사유로 위장전입이 많고 해외 유학이나 취업 등 지연 신고나 누락이 발생한다(김민경, 2002).

하지만 주민등록인구는 병무, 취학, 선거, 국고보조금, 지방행정기관 직제 등 지방행정기관의 행정업무와 밀접하게 관련이 되어 있어, 주민등록인구는 실제로 센서스 실시 과정에서 지방행정기관의 심리적 가이드라인이 되고 있다. 즉, 상주개념에 의해 센서스를 실시하면서도 관할 구역의 상주인구가 주민등록인구 수준과 비슷해야 한다는 심리적인 압박감을 갖게 된다. 이 점은 사실 조사현장에서의 실사지도 태도나 행태에 커다란 영향을 미치고 있다. 지방 공무원, 심지어 통계청의 센서스 담당자조차도 센서스인구가 주민등록인구와 커다란 괴리를 보이면 당혹하지 않을 수 없게 된다.

2. 유럽 및 미국의 인구센서스와 행정자료

행정자료를 센서스에 활용하는데 있어 선도적인 나라는 유럽의 국가들이다. 이들은 중복응답에 대한 거부감, 사생활 침해, 조사비용 절감이라는 이유로 전통적인 센서스 방식에서 탈피하여 행정자료에 눈을 돌리게 되었다. 사실 유럽의 센서스는 많은 도전에 직면해 왔다. 독일, 네덜란드에서는 국민들의 반발로 센서스를 실시하지 못하고 있는 형편이다. 1971년 네덜란드 센서스에서는 사생활 침해를 이유로 30만명(인구의 2%이상)이 응답을 거부했는데, 1979년 센서스 참여 실험에서는 무응답률이 25%에 달한 가운데 주요 도시에서는 40~50%의 무응답률을 보였다. 1990년 스웨덴 센서스의 경우에도 스톡홀름에서만 무응답이 10%에 달했다.

이에 따라 노르딕 국가를 중심으로 전통적인 센서스 방식을 벗어나 인구등록자료로 센서스를 대체하려는 움직임이 시작되었다. 덴마크, 핀란드, 아이슬란드, 노르웨이, 스웨덴은 1960년대 경부터 지방주민등록제도에 의거해 인구등록제도를 도입했다. 이들 국가에서 센서스의 인구부문을 위한 행정자료는 1970년대에 이미 작성이 가능했다. 전반적인 행정자료 기반의 센서스가 완성된 시기를 보면, 덴마크는 1981년, 핀란드는 1990년이였다. 스웨덴은 행정자료 베이스의 고용통계를 1985년에 도입하였지만, 거처와 건물 등록은 아직 하지 못했다. 노르웨이는 스웨덴과 비슷한 단계인데, 1990년에 실시한 표본 센서스 때문에 오히려 행정자료의 활용이 더디게 되었다. 아이슬랜드는 행정자료에 인구자료와 주택자료

는 있지만 고용자료는 부족하다. 덴마크, 핀란드, 아이슬랜드의 2001년 센서스는 전적으로 행정자료에 기반을 둔 센서스이다. 반면, 노르웨이와 스웨덴의 센서스는 인구나 경제활동은 행정자료를 이용하지만 주택사항은 센서스 조사표로 조사한다.

<표 1> 유럽의 인구센서스 실시 형태

센서스 형태	국가명	실시 내용
전부 또는 대부분 행정자료 기반의 센서스	덴마크	- 세계 최초로 행정자료만을 이용한 센서스 실시(1981)
	핀란드	- 세계 두 번째로 행정자료만을 이용한 센서스 실시(1990)
	노르웨이	- 1990년 표본 센서스 실험 - 2001년 센서스에서 인구 및 경제활동은 행정자료로, 주택사항은 조사표로 조사
	스웨덴	- 1985년 경제활동을 행정자료로 작성 - 1990년 센서스는 기존방식에 의한 마지막 센서스 - 2001년 센서스는 인구 및 경제활동은 행정자료로, 주택사항은 조사표로 조사
전통적인 방식에서 행정자료 기반의 센서스로 전환 중	오스트리아	- 2001년 센서스는 전통적인 방식으로 했지만, 건물등기부는 현장조사 준비에, 사업체 및 농업 등록명부는 대상처 확인 및 코딩에 사용하고, 사업체 규모 및 법적지위는 행정자료로 대체
	벨기에	- 인구 및 사회자료는 인구등록 및 사회보장 파일을 이용하되, 부족한 자료는 25%표본으로 조사
	룩셈부르크	- 2001년을 반(半) 전통방식의 센서스로 실시 (등록자료의 질이 취약)
	스위스	- 조사표로 조사하되 행정자료를 활용 - 제5차 조사가 마지막 전통방식의 센서스로 예상
다른 대안을 모색 중인 센서스	프랑스	- 전통적인 방식으로 조사하되, 행정자료 활용(거처등록자료는 자료 수집과정에서, 사업체 등록자료는 산업부호 코딩에 활용) - 순환조사(5년 간격)로 지속적인(continuous) 센서스를 구상 중
	독일	- 2001년 센서스 자료를 주민등록 기초의 인구추계로 작성 - 소지역 고용통계는 고용등록자료, 1% 마이크로 센서스나 다른 자료원을 이용
	네덜란드	- 행정자료와 표본조사를 통합하여 센서스 자료를 작성 - 행정자료 이용 수준은 스칸디나비아 국가와 비슷하게 광범위
전통적인 센서스	이탈리아	- 2001년에 센서스를 실시했으며, 처음으로 빌딩센서스와 함께 실시 - 행정자료는 지도제작과 현장작업에만 제한적으로 활용
	그리스, 포르투갈	- 2001년에 센서스를 실시했으며, 행정자료는 지도제작과 현장조사 지원 등에만 활용
	스페인	- 2001년에 센서스를 실시했으며, 법 개정으로 5년마다 하던 주민등록 갱신을 상시적으로 실시하여 장차 센서스와 통합을 기대
	아일랜드	- 2001년에 센서스를 실시했으며, 행정자료의 활용은 미미
	영국	- 2001년에 센서스를 실시했으며, 행정자료 활용은 미미하여 현장조사만 지원

자료: Eurostat(1999).

전체적으로 볼 때, 덴마크와 핀란드가 세계에서 가장 성공적으로 인구등록자료를 센서스 실시예 활용한 국가로 나타났다. <표 1>을 보면, 이제 유럽에서 전통적인 센서스를 실시하는 국가는 이탈리아, 그리스, 포르투갈, 스페인, 아일랜드, 영국 정도에 불과하다. 덴마크, 핀란드를 비롯해 노르웨이와 스웨덴이 전부 또는 대부분 행정자료를 기반으로 하는 센서스를 실시하고 있으며, 오스트리아, 벨기에, 룩셈부르크, 스위스는 전통적인 방식에서 행정자료 기반의 센서스로 전환 중에 있다. 물론 프랑스, 독일, 네덜란드에서도 마이크로 센서스(micro-census) 또는 표본 센서스(sample-census), 순환 센서스(rotation-census) 등 다른 대안들을 검토하는 등 이제 전통적인 센서스 방식이 커다란 변화에 직면해 있는 것만은 틀림이 없다.

미국에서는 행정자료를 본격적으로 센서스에 활용한 것이 1980년 센서스부터인데, 주로 조사율(coverage)이나 주소록 개선 등의 목적으로 쓰여 지고 있다. 예를 들어, 자동차 등록, 주택, 가석방이나 집행유예자 명부 등이 사회보장번호(Social Security Number: 이후 SSN으로 표기)⁵⁾와 주소로 센서스 자료와 연계된 후 누락을 보완하는데 사용된다. 우편조사(mail-out/mail-back)를 위한 주소는 미국 우편서비스(U.S. Postal Service)의 주소록을 사용하는 데, 소득세 화일로 주소록 개선을 도모하고 있다. 1995년 센서스 테스트(Paterson, Oakland, Louisiana 등 3개 지역에서 실시)에서는 처음으로 센서스 자료와 다양한 행정자료⁶⁾의 연결 및 통합을 시도했는데, 개인자료 연결은 SSN을 이용하여 주소에서 시작하여 지역으로 확대해 나갔다. 주소와 관련된 문제⁷⁾가 많아 센서스와 행정자료가 Paterson은 11%, Oakland는 25%, Louisiana는 21%만이 연계되었다. 이로써 미국 센서스국은 행정자료 파일의 접근성, 내용, 연결방법 등에서 많은 경험을 갖게 되었다.

미국에서 행정자료 기반의 센서스 개념이 제기된 것은 2000년 센서스부터이지만, 1990년 센서스에서 행정자료에 대한 테스트를 제대로 하지 않아 심층적으로 문제를 파악하고 해결할 시간이 없다는 이유로 2000년 센서스에서는 행정자료의 활용이 오히려 위축되었다. 하지만 2000년 센서스에서는 다양한 실험을 포함하고 있어, 2010년 센서스에서는 보다 광범위한 행정자료 활용이 예상되고 있다.

5) 하이픈으로 3자리(area number)-2자리(group number)-4자리(serial number)로 구분하여 총 9자리로 구성되어 있다.

6) 처음에는 40개의 상업용, 연방·주·지방정부의 파일을 이용해 통합을 시도했으나 필요 없는 정보가 많아 예산과 시간을 많이 허비하게 되자, 다음과 같은 전국단위 자료로만 한정했다; 납세자료(Internal Review Service Tax Return File), 주거 및 도시개발성 자료(Department of Housing and Urban Development File), 군복무 자료(Selective Service System File), 인디언 보건서비스 자료(Indian Health Service File), 사회보장청 자료(Social Security Administration File) 등

7) 오래된 주소, 거주지 주소대신 사서함 우편주소가 등재된 경우, 화일간 주소가 다른 경우 등

Ⅲ. 인구센서스와 행정자료의 통합모형 제시

이 장에서는 우리나라에서 인구센서스와 행정자료를 통합할 때 고려할 수 있는 절차나 모형을 제시하고자 한다. 인구센서스와 행정자료를 통합하기 위해서는 먼저 행정자료에 대한 실태를 파악하는 것이 중요하다. 이용 가능한 행정자료는 어떠한 것이 있으며, 자료의 작성기관, 주기, 개념 및 포괄범위, 제약성, 정보의 질 및 활용가치 등에 대한 세부적인 실태 파악이 선행되어야 한다. 참고로 주요 행정자료의 종류와 수록내용은 <부록 2>~<부록 4>에 제시하였다.

다음 단계로는 인구센서스와 행정자료를 통합할 수 있는 기술적인 방법이 모색되어야 한다. 여기서는 <그림 1>과 같은 모형으로 접근하고자 한다. 우리가 통계자료를 집계단위로 구분하면 크게 인구단위의 자료(P1, P2, P3…), 가구단위의 자료(H1, H2, H3…) 및 사업체 단위의 자료(E1, E2, E3…)로 나눌 수 있다. 이 단위별 자료는 인구DB군, 가구DB군, 사업체 DB군으로 묶어지게 된다.

여기서 동일한 DB군의 자료들은 서로 단위가 같기 때문에 1차적으로 통합이 가능하다. 이 때 이들을 서로 연결하기 위해서는 공통으로 사용할 수 있는 인구 key, 가구 key, 사업체 key가 각각 필요하다. 예를 들어 인구 key는 주민등록번호, 가구 key는 가구번호, 사업체 key는 사업체번호와 같이 서로 공유할 수 있는 ID를 사용해야 한다.

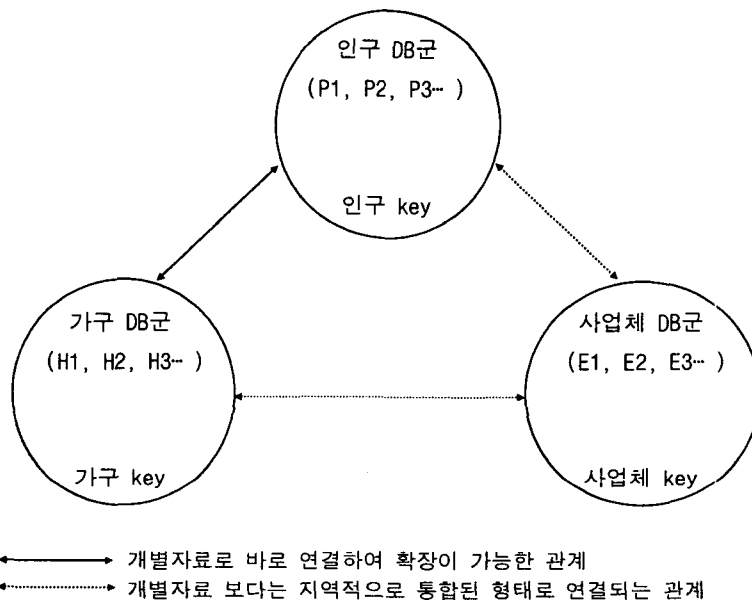
서로 다른 DB군의 결합은 인구 DB군과 가구 DB군에서만 직접 연결이 가능하다(<그림 1> 참조). 예를 들면, 인구 DB군 자료와 가구 DB군 자료는, 물론 개념상에 논란의 여지가 없는 것은 아니지만, 가구주(또는 세대주, 대표자)라는 변수를 이용하면 서로 연결이 가능하여 개별 자료별로 정보를 확장할 수 있다. 이 두 가지 DB군의 통합은 주민등록번호, 이름은 물론 주소도 보조수단이 될 수 있다.

반면 <그림 1>에서 나타나듯이 다른 DB군 간에는 일부 제한적인 범위에서만 개별 자료끼리 연결이 가능하다. 이 자료들은 대부분 지역적으로 통합된 자료(spatially aggregated data)의 형태에서만 연계될 수 있다. 여기서 통합의 스케일(scale of aggregation)은 기존의 지역구분보다 훨씬 자유롭고 다양할 수 있다. 예를 들어 이 DB군들이 읍·면·동 보다 작은 수준의 조사구, 기초단위구 또는 반경 500m, 반경 1km 등의 스케일로도 통합되어 다양한 소지역통계를 작성할 수 있다. 이 때 DB군 간의 지역적 통합은 개별 자료의 절대좌표 값을 제공할 수 있는 지리정보시스템의 지원으로 보다 효과적이고 강력해 질 수 있을 것이다.

자료통합을 위한 틀이 만들어지면, 법과 제도적 장치에 대한 검토에 들어가야 한다. 현재의 법체계로는 행정자료에 수록되어 있는 개인정보에 관한 사항은 통계목적으로 바로 활용할 수가 없어, 통계목적으로는 원시자료를 제한적으로 사용할 수 있도록 행정관련 개별법령의 보완이 필요하다. 이러한 개별법령의 정비에는 각 부처에서 개별 자료의 제공을

꺼리고 있는데다가, 이런 이유로 법 개정을 요구하기가 힘들어 다른 이유로 법률을 개정할 때까지 기다려야 하는 등 실질적으로 많은 시간이 소요되는 힘든 일 중의 하나이다. 따라서 법령 개정은 중장기적으로 추진하되, 통계청 소관법령인 통계법부터 개정해 착수해야 할 것이다.

<그림 1> 인구센서스 자료와 행정자료의 통합 모형



주) P1=인구센서스, P2=주민등록, P3=호적, P4=소득세, P5=국민연금, P6=건강보험, P7=고용보험, P8=산재보험, P9=외국인등록... ; H1=인구센서스, H2=토지등기, H3=건물등기... ; E1=사업자등록, E2=사업체기초조사, E3=협회 자료...

다음은 분류 시스템의 체계화 및 통일 작업이 필요하다. 먼저 주소체계의 경우 현재의 지번에 바탕을 둔 주소체계로는 가구나 주택을 구별하기가 어렵다.⁸⁾ 예를 들어 1개 번지에 다수의 가구가 속해 배타적인 가구 주소가 없다. 이를 위해 정부에서는 1996년부터 모든 도로에 명칭을 부여하고 도로에 면한 건물에 건물번호를 부여하는 “새주소 부여 사업”을 전국적으로 실시하고 있는 바, 행정자료를 통계자료에 통합하여 활용하려는 입장에서 보면 크게 다행스러운 일이다. 하지만 건물별로 부여되고 있는 새 주소가 여러 개의 건물

8) 우리나라의 주소체계는 1910년에 설정된 지번체계에 기초하고 있으나, 처음 설정 후 많은 변화를 거쳐 지번 배열순서가 불합리하고 상대적 위치식별이라는 측면에서 거의 그 기능을 발휘하지 못하고 있다(서울특별시, 2002).

이 하나의 건물로 합쳐질 경우 관리기능이 제대로 작동하지 않는다거나, 기존의 행정자료 주소와 어떻게 성공적으로 연계시킬 수 있는가와 같은 일부 의문점에 대해서는 지속적인 개선이 필요한 것으로 보인다.

산업분류나 직업분류도 통계청에서 표준분류를 제공하고 있으나 통계이외의 다른 목적으로는 부적합한 경우도 있어 사업체나 보험 등에서 일부 이를 달리 사용하기도 한다. 만약에 자료 통합이 이루어진다면, 이에 대해서도 좀더 세밀한 검토가 필요하다.

행정 및 통계자료의 통합시스템은 정부 각 부처 및 지자체, 연구기관, 공단, 협회, 기업 등 DB 자료를 보유하고 있는 다양한 기관이 관여하기 때문에, 통계청과 같은 중립적인 기관이 전체 데이터 웨어하우스를 관리하고 운영할 수 있도록 해야 한다. 개별 DB를 보유하고 있는 기관에서는 제한적으로만 다른 DB에 접근하도록 하고 있는 북유럽 국가의 예도 좋을 것 같다.

행정자료와 조사통계가 통합되면 우리 사회는 지금보다 더 많은 정보를 저비용으로 축적할 수 있다. 국민의 응답부담도 크게 줄어들게 된다. 우선 인구센서스의 경우 전수조사 대신에 필요한 부분만 표본조사로 실시할 수도 있는데, 내용은 행정자료와의 통합으로 더욱 방대해 지는 반면 비용은 대폭 줄어들게 된다. 유럽에서 행정자료 기반의 센서스 전환에 따른 경제적 효과를 보면, 스웨덴은 2005년 센서스 비용이 200만 달러가 들것으로 보여 1985년(2,500만 달러)의 1/12 수준으로 떨어지며, 핀란드는 1991~2000년(10년간) 800만 유로가 들어 1980년(3,500만 유로)에 비해 자료의 양은 10배 늘면서 비용은 1/4로 줄어든 것으로 나타났다.

자료면에서 보면 센서스와 같은 방대한 내용을 매년 또는 더 짧은 주기로도 작성할 수 있어 자료의 양이나 시의성에서 획기적인 개선을 가져오게 된다. 이러한 자료들이 시계열로 축적이 되면 바로 패널 데이터로 이용할 수 있어 인구센서스의 패러다임이 크게 변하게 된다.

그런데 이러한 모형 구축과정에서 반드시 거쳐야 할 사항중 하나는 개인정보의 통계목적 활용에 대한 국민들의 이해와 협조가 필요하다는 것이다. 아무리 행정자료와 조사통계의 통합이 장점이 많더라도 국민이 반대한다면 현실적으로 추진이 불가능하다.

IV. 인구센서스와 주민등록의 활용 및 통합

이 장에서는 인구센서스와 행정자료의 통합 모형의 근간을 이루는 인구센서스와 주민등록에 대해 양 자료의 차이점과 상호 연계 가능성을 살펴본다.

1. 인구센서스 자료와 주민등록 자료의 차이

인구센서스는 조사 시 3개월 이상 거주하고 있거나 거주할 목적으로 살고 있는 상주(de jure)인구를 조사한다. 주민등록과 관계없이 실질적으로 거주하는 장소에서 조사하기 때문에, 주민등록은 되어 있는데 다른 곳(양로원, 기도원, 군대, 교도소, 선박, 무단가출 등)에서 살고 있는 사람은 그 곳에서 조사한다. 이에 따라 3개월 이상 해외로 취업, 유학, 연수를 나간 사람은 조사하지 않게 된다(통계청, 2000).

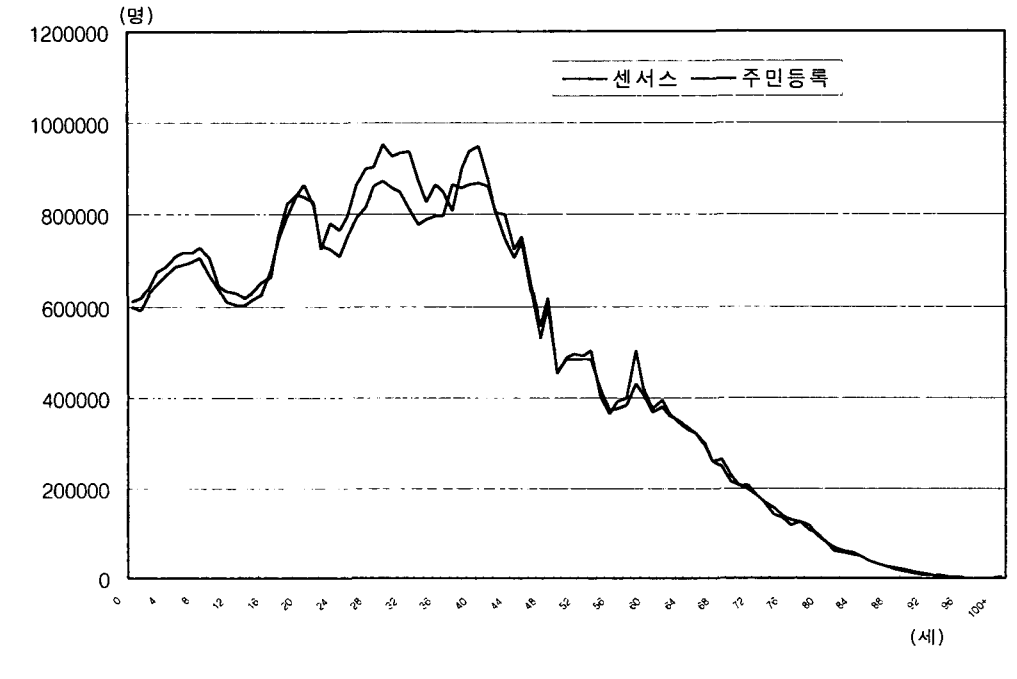
반면에 주민등록은 법에 의거 1개월 이상 거주할 목적으로 주소 또는 거소를 옮기면 14일내에 관할 시장, 군수 또는 구청장에게 기본적인 인적사항을 신고하도록 되어 있다. 하지만 주민등록은 행정목적의 등록자료이기 때문에 사실을 조사하는 조사통계와는 달리 신고자의 신고내용이 사실과 다른 경우에도 검증 없이 등재되고 있다. 사실과 다른 예를 보면, 주민등록을 실제 거주지가 아닌 지역에 등재하거나, 실제로 동일한 가구원이면서도 주민등록상으로는 별도 세대로 등재하거나, 출생년월일이 제대로 신고 되지 않거나, 출생·사망·혼인·이혼 등 변동사항에 대한 지연신고 및 누락, 해외유학생, 해외취업자, 이민자, 주민등록 말소자 등에 대한 주민등록 미등록 등으로 나타나고 있다.

주민등록자료 사용시 나타날 수 있는 문제점으로는, 지역별(특히, 소지역) 인구의 규모, 성 및 연령별 특성, 분포 및 이동에 관한 현상파악이 왜곡될 수 있어 자칫 자원배분의 비효율 및 낭비의 개연성이 있다는 것이다. 또한 주민등록상의 세대는 사회학적인 생활단위인 가구와는 개념이 달라 가구를 기초로 파악해야 하는 각종 경제·사회 지표의 작성이 곤란할뿐더러 가족구조의 변동이나 탈가족화 등 관련문제에 대한 심층분석이 어렵다.

이처럼 조사개념 및 포괄범위의 차이로 두 인구자료는 차이를 보이게 된다. 해외 유학생 및 해외 취업자, 어학연수생 등과 같은 포괄범위의 차이에도 센서스의 일부 조사 누락 때문에 통상 주민등록인구가 센서스 인구보다 많다. <그림 2>에서는 두 자료간의 기준시점차이(센서스는 2000.11.1일 기준, 주민등록은 2000.12.31일 기준으로 주민등록이 2개월 늦어 인구가 더 많은 요인 중 하나가 됨)를 보정하지 않고 연령별로 대비시켜 보았다. 전체적인 패턴을 보면 두 자료 간에 차이가 거의 없어 양 자료가 모두 인구통계의 자료원으로 손색이 없음을 보여 주고 있다. 물론 20대 후반에서 30대에 걸쳐 주민등록 인구가 센서스보다 많이 나타나는데, 이는 개념상의 차이에도 이들 연령이 이동성향이 강하고 1인 가구 비율이 높아 조사측면에서도 취약계층이기 때문이다. 주민등록자료의 연령이 신고 지연 때문에 하향 보고 되어 차이가 많을 거라는 우려와는 달리 최근에는 실제 연령과 상당히 근접해 가고 있다.

여기서 제시하지는 않지만 지역별(읍·면·동 수준)로 볼 때는 특·광역시 등의 경우에는 센서스 인구보다 주민등록 인구가 많은 지역이 많았으며, 특·광역시 인접 동이나 일반 시 및 군의 경우에는 반대로 주민등록인구보다 센서스 인구가 많은 지역이 많았다.

<그림 2> 연령별 센서스인구(2000.11.1) 및 주민등록인구(2000.12.13)의 비교



한편, 센서스 자료와 주민등록 자료를 보다 잘 연계, 활용하기 위해서는 양 자료의 개념 및 포괄범위에 대한 조화가 필요하다. 자료의 질적인 측면에서 보면, 센서스는 현장조사의 정확도가 제고되어야하고 주민등록은 신고의 질이 더욱 개선되어야 한다. 하지만 두 자료 간의 이러한 차이에도 불구하고 통합에 대한 매력이나 장점이 줄어드는 것은 아니다.

2. 주민등록번호 조사와 개인정보보호의 문제

앞에서 인구센서스와 주민등록자료를 연결할 수 있는 수단으로 인구 Key로서의 주민등록번호에 대해 논의한 바 있다. 주민등록번호를 연결 key로 사용하기 위해서는 인구센서스에서 주민등록번호를 조사해야 하지만, 어려움이 많을 것이다. 조사 불응과 조사과정에서 알게 되는 주민등록번호의 악용 문제를 간과할 수 없기 때문이다.

현재 미국 센서스국에서도 비슷한 고민을 하고 있다. 미국에서도 간결한 센서스 조사표로 응답부담을 경감하여 응답의 완전성을 증진하고, 기억 및 응답오차 제거로 질을 향상시키기 위해 행정자료의 이용 가능성을 타진하고 있다. 2010년 센서스는 다른 연방기관의 행정자료에 보다 많이 의존할 예정이다(Neugebauer, 1999).

이와 관련하여 인구센서스와 행정자료의 연계수단으로 조사표에서 개인의 SSN을 조사하는 방안에 대해서 꾸준히 타당성을 검토해 왔다. 센서스 조사표에서 SSN을 조사하는데 대한 대중의 의견과 반응행태(우편회수율, 조사표 완전기입률, SSN 항목 응답률, 조사된 SSN의 유효성)를 측정된 결과를 보면, 고응답층 보다는 저응답층에서 SSN 질문의 응답률이 낮았으며(Guarino et al., 2001), 응답자가 개인 식별이 가능한 정보의 제공에 반대할 뿐더러 다른 가구원의 정보를 얻는데 어려움이 있어 SSN 질문이 응답의 완전율을 낮추는 것으로 나타났다(Dillman, Sinclair, and Clark, 1993). 또한 Singer 외(2001)에 따르면, 센서스 조사표상에 SSN 제공용의 비율은 1996년 68%, 1999년 55%, 2000년 56%로 각각 나타났다. Bates(1992)는 1번째 이후 가구원부터는 SSN 번호를 얻기 어렵다는 연구결과를 내놓았는데, 응답률이 낮은 것이 용의가 낮은 것이 아니라 때로는 다른 가구원(특히 아동 및 비혈연가구원)의 SSN에 접근하기 곤란한 것 때문으로 파악되었다. 다른 연구(Dillman, Sinclair, and Clark, 1993)에서도 SSN을 거의 사용하지 않는 아동의 응답률이 낮았는데, 17세미만 아동은 무응답 비율이 25%에 달했다.

미국 센서스국에서 2000년 센서스 테스트, 실험 및 평가 프로그램으로 실시한 SPAN(The Social Security Number, Privacy Attitudes, and Notification) 실험에서는 조사표를 통해 SSN을 조사하고, 조사표에 이를 이용해 다른 행정자료를 사용할 것임을 통지⁹⁾해 주었을 때, SSN의 응답행태가 어떻게 변하는지를 테스트했다. 이 연구에서는 조사된 SSN이 비교적 정확한 것으로 나타났으며, 응답률이 높은 지역(HCAs)과 낮은 지역(LCAs)간에 차이는 적었지만 통계적으로는 유의한 것으로 나타났다. 1번 가구원만 SSN을 묻는다면, 모든 가구원의 SSN을 묻는다면, 일반적인 통지를 하든, 구체적인 통지를 하든 SSN의 결과에 영향을 미친다는 증거는 없었다(Brudvig, 2003).

센서스에서 주민등록번호를 조사하는 문제에 대해서는 미리 겁을 집어 먹을 필요도 없지만, 국민들의 개인정보보호와 밀접하게 관련이 되어 있으므로 아무리 신중해도 지나침이 없다. 주민등록번호를 응답자에게 직접 조사하는 대신, 센서스 조사항목에서 출생년월일 및 세는 나이를 주민등록상의 생년월일로 바꾸어 조사하면 주민등록번호와 연결할 수 있는 대략적인 정보를 얻을 수는 있지만, 이와 같은 방법 또한 국민들의 절대적인 이해와 협조가 필수적이다.

9) 통지는 일반적인 통지와 구체적인 통지의 2가지로 구분해 조사표에 표시되었다. 일반적인 통지는 “센서스 자료의 질 개선을 위하여 센서스국에서는 가끔 다른 정부 기관의 행정자료를 이용합니다. 다른 행정자료를 이용하는 것은 센서스를 보다 정확하게 만듭니다. 현존하는 정부 기록을 보다 잘 이용하기 위해서 센서스국에서 몇 가지 더 질문할 수도 있습니다.”였으며, 구체적 통지는 “센서스 자료의 질 개선을 위하여 센서스국에서는 가끔 예를 들면, 사회보장청(the Social Security Administration)이나 국세청, 또는 국민 주거지원을 해주는 기관과 같은 다른 정부 기관의 행정자료를 이용합니다. 다른 행정자료를 이용하는 것은 센서스를 보다 정확하게 만듭니다. 현존하는 정부 기록을 보다 잘 이용하기 위해서 센서스국에서 몇 가지 더 질문할 수도 있습니다.”라고 조사표에 표기되었다(Brudvig, 2003).

V. 결론 및 토론

이 글에서는 인구센서스의 비용증가와 국민의 조사 비협조에 대한 대안으로 행정자료의 활용 내지는 나아가 인구센서스와 행정자료의 통합 필요성과 방안에 대해 살펴보았다.

먼저 한국의 인구센서스에서의 행정자료 활용에 대해 살펴보면, 그동안 행정자료가 조사구 설정, 지도제작, 조사개념 및 자료분석을 위한 참고자료로 주로 사용되어 왔는데, 2000년 센서스에서는 진일보하여 “공동주택 기본정보 사전수집 및 활용” 등 좀 더 적극적으로 행정자료를 현장조사에까지 활용하고 있는 것으로 나타났다. 하지만 이제는 각종 행정자료들이 IT 기술의 발전으로 자료의 축척, 접근 및 활용이 용이해지고 있어, 인구센서스와 행정자료를 개별수준에서 연계할 필요성이 크게 제기되고 있다.

행정자료를 센서스에 활용하는 점에서 선도적인 지역은 유럽의 국가들이다. 유럽에서는 조사비용과 사생활보호 또는 중복응답에 대한 국민들의 반감으로 센서스 실시에 어려움을 겪다가, 행정자료로 눈을 돌려 행정자료 기반의 센서스를 실시하기에 이르렀다. 세계에서 처음으로 덴마크(1981년)가 행정자료 기반 센서스로 성공적으로 전환했고, 핀란드(1990년)도 전환을 마쳤다. 행정자료 기반의 센서스에 따른 경제적인 효과를 보면 스웨덴은 센서스 비용이 전통적인 조사의 1/12, 핀란드는 자료는 10배가 많아졌는데 비용은 1/4만 소요 되는 것으로 나타났다.

하지만 유럽 국가들이 성공적으로 행정자료를 이용한 센서스를 실시할 수 있는 배경에 대해서도 논의할 필요가 있다. 이 들 국가들이 일찍이 출생, 사망 등 인구동태를 교구(parish)에 등록하기 시작(스웨덴 1571년, 덴마크 1743년)하여 비교적 인구등록 시스템이 잘 구비되어 있었다는 점이 충분히 감안되어야 한다. 게다가 거주지 등록이 연금, 의료보험, 승차권 등 복지제도와 밀접하게 연계되어 있고, 법률도 거주지 이동시 1주일 내 신고를 의무화하고 있다. 즉, 북유럽에서 인구등록이 성공을 거둔 데는 거주지 등록의 오랜 전통과 지역기반의 사회적 혜택 그리고 잘 정비된 법률의 역할이 컸다. 아울러 이들 국가가 행정자료 기반 센서스를 실시하기까지는 스웨덴이 30년(1975~2005년), 핀란드가 15년(1985~2000년), 덴마크가 11년(1970~1981년) 등 약 10~30년간의 시간을 투자하였음을 주목해야 한다.

미국에서도 행정자료의 센서스 이용이 활발하게 검토되고 있다. 현재로는 조사율(coverage)이나 주소록 개선 등에 제한적으로 활용되고 있지만, 1995년 센서스 테스트부터 센서스 자료와 다양한 행정자료의 접목을 시도하고 있어 경험이 상당히 축적되어 있다. 특히 인프라면에서 보면, 이미 1970년대부터 SSN을 이용해 소득세 자료와 SSN 화일을 연계하여 사용할 수 있는 법적인 지위를 갖고 있어, 행정자료 기반 센서스 실시를 위한 토대가 탄탄하게 구축되어 있다. 2000년 센서스에서는 1990년 센서스에서 행정자료에 대한 테스트를 제대로 하지 않았다는 이유로 일시적으로 위축되어 있기는 하지만, 2000년

센서스에서 다양한 실험을 포함하고 있어 2010년 센서스에서는 보다 광범위한 행정자료 활용이 예상되고 있다.

한편, 이 연구에서는 인구센서스와 행정자료의 통합과정과 모형이 제시되었다. 먼저 행정자료에 대한 실태파악이 선행되어야 함을 지적했는데, 행정자료의 종류도 중요하지만 각 자료들이 가지고 있는 개념, 포괄범위, 자료의 제약성, 자료의 질 등 실질적인 내용 파악이 중요한 것으로 논의되었다. 통합모형은 크게 인구 DB군, 가구 DB군, 사업체 DB군으로 나누어지는데, 각 DB군내에서는 각각의 Key로 자료들을 통합하게 된다. DB군 간에는 인구DB와 가구DB는 개별 단위로, 나머지는 지역별로 통합된 수준에서의 연계를 제시하였다. 또한 법과 제도 측면에서도 행정자료를 활용할 수 있도록 정비가 필요한 것으로 지적되었는데, 행정여건을 보면 이러한 작업들은 중장기적으로 이루어져야 할 것으로 보인다. 또한 개인정보의 통계목적 활용에 대한 국민들의 이해와 협조의 중요성도 강조되었다. 이와 관련해 정부에서는 먼저 국민들의 프라이버시나 개인정보의 통계목적 활용에 대한 태도에 대해 객관적으로 측정하고, 이 결과를 개인정보의 안전장치와 통합의 필요성을 홍보하는데 반영해야 할 것이다.

마지막으로 통합 모형 중 가장 근간을 이루는 인구센서스 자료와 주민등록자료에 대한 비교 검토에서는, 인구통계로서 양 자료가 전반적으로 동일한 패턴을 보이는 것으로 나타났다. 물론 연령별로는 이동성향이 강한 20대 후반에서 30대에서 차이를 보였으며, 지역적으로는 특·광역시는 주민등록인구가, 그 주변지역이나 일반 시와 군에서는 인구센서스 인구가 상대적으로 더 많은 지역이 많이 나타나, 향후 두 자료의 연계시 보완해 나갈 점으로 지적되었다.

양 자료를 연계할 수 있는 주요 수단인 주민등록번호를 인구센서스에서 조사하는 것과 관련해서는, 센서스 조사표를 이용해 SSN을 조사하는데 따른 미국 국민들의 태도조사 실시 결과가 우리에게 많은 것을 시사해 주고 있다. 여기서는 센서스에서 조사된 SSN이 비교적 정확했으며, 아동이나 비혈연가구의 SSN은 무응답이 높았지만 전체적인 정확도에는 유의하게 영향을 미치지 않는 것으로 나타났다.

우리 사회는 인구센서스와 주민등록이라는 엄청난 인구통계의 인프라를 갖고 있다. 이제 국민의 통계 응답부담을 경감하고, 국가재원을 보다 효율적으로 사용한다는 측면에서 이 두 개의 커다란 데이터 웨어하우스는 얼마가 걸릴지는 모르지만 통합의 길을 모색해야만 한다.

참고문헌

- 김민경. 2000. 『인구센서스의 이해』. 도서출판 글로벌.
- _____. 2002. “인구센서스의 발전과 특징.” 김두섭·박상태·은기수 편. 『한국의 인구』. 통계청.
- 이삼식·박종서·하미영. 2002. 『환경변화에 따른 인구주택총조사 개선방안』. 통계청·한국보건사회연구원.
- 서울특별시. 2002. 『서울시 도시계획 정보관리시스템 유지관리방안 연구』.
- 통계청. 2000. 『2000 인구주택총조사 지도지침서』.
- _____. 2002. 『인구센서스의 개선방안에 관한 연구』.
- _____. 2003. “한국통계정보시스템(KOSIS).” <http://kosis.nso.go.kr>.
- Bates, N. 1992. “Revised Item Nonresponse Results for Social Security Number From the Simplified Questionnaire Test(SQT).” *Memorandum to Robert D. Tortura and Susan M. Miskura*.
- Brudvig, Linda. 2003. “Analysis of the Social Security Number-Validation Component of Social Security Number, Privacy Attitudes, and Notification Experiment.” *Census 2000 Testing, Experimentation, and Evaluation Program*. U.S. Census Bureau.
- Dillman, Don A., Michael D. Sinclair, and John R. Clark. 1993. “Effects of Questionnaire Length, Respondent Friendly Design, and a Difficult Question on Response Rates for Occupant-Addressed Census Mail Surveys.” *Public Opinion Quarterly* 57: 289-304.
- Eurostat. 1999. “Development of the Use of Administrative Data in Population and Housing Censuses in Europe.” *An Invited Paper to the Joint ECE/Eurostat Work Session on Registers and Administrative Records for Social and Demographic Statistics* March 1-3. Geneva.
- Guarino, Jennifer A., Joan M. Hill, and Henry F. Woltman. 1999. “Census 2000 Experimentation Program Master Plan: The Social Security Number, Privacy Attitudes, and Notification Experiment.” U.S. Census Bureau.
- Kim, Hyung-seog. 2001. “Who Should We Focus on to Ensure a Better Count: The experience of the 2000 Population and Housing Census in Korea.” *An Invited Paper to the 53rd Session of the International Statistical Institute* August 22-29: Seoul, Korea.
- Neugebauer, Randall. 1999. “Census 2000 Experimentation Program Master Plan: The

Social Security Number, Privacy Attitudes, and Notification Experiment.” U.S. Census Bureau

Singer, E., J. Van Hoewyk, R. Tourangeau, D.M. Steiger, M. Montgomery, R. Montgomery. 2001. “1999-2000 Surveys of Privacy Attitudes.” U.S. Census Bureau

Suliman, Sirageldin H. 1998. “Use of Administrative Records in Population Censuses and in Other Demographic and Social Statistics”. UN Statistics Division Home Page.

U.S. Census Bureau. 1999. “Uses of Administrative Records in United States Census 2000.” *An Invited Paper to the Joint ECE/Eurostat Work Session on Registers and Administrative Records for Social and Demographic Statistics* March 1-3. Geneva.

<부록 1> 행정자료의 통계작성 활용 예(통계청)

통계조사	활용 행정자료	활용분야	자료수준
- 건설업 및 운수업 통계조사	· 협회 등록자료	· 명부	· 개별자료
- 광공업통계조사	· 금감원 상장회사 재무회계 자료	· 내용검토시	· 개별자료
- 농어업 법인사업체 조사	· 법인등기대장 · 시·군·구 등록자료 · 각종 협회자료	· 표본틀	· 개별자료
- 지역내 총생산	· 27종의 행정자료	· 통계작성	· 지역자료(시·도)
- 사망원인	· 암등록 자료(국립암센터) · 건강보험자료(공단) · 산재보험(공단) · 법정전염병(국립보건원)	· 사망원인 보완(신고자료가 부정확할 때)	· 개별자료
	· 모자보건(보건소) · 화장장 자료(화장장)	· 영아사망 보완	· 개별자료
- 인구이동	· 주민등록	· 통계작성	· 개별자료(생년월일, 성별, 주소)

<부록 2> 행정정보의 공동이용 시스템(G4C) 주요 수록정보

행정자료 종류	제공기관 (담당부서)	수록 내용
주민등록	행정자치부 (주민과)	성명, 세대주, 주민등록번호, 세대주와의 관계, 호주소항, 세대원수 등
건물등기	대법원 (법정심의관실)	지번, 면적, 용도, 소유자의 성명, 주민등록번호, 주소 등
토지등기	대법원 (법정심의관실)	지번, 지목(임야), 면적, 소유자 성명, 주소 등
호적자료	대법원 (법정심의관실)	성명, 본적지, 호주와 전 호주와의 관계, 본관, 주민등록번호, 출생지, 혼인신고일 등
자동차 및 이륜자동차	건설교통부 (자동차관리과)	등록번호, 형식승인번호, 차명, 차종, 차대번호, 원동기형식, 용도, 년식, 색상, 출처구분, 최초등록일, 최초접수번호, 제작년월일, 최종소유자, 주민등록번호, 사용본거지(차고지) 등
납세 관련 행정자료	국세청 (법인세/소득세)	성명(대표자), 주민등록번호, 상호(법인명), 사업자등록번호, 주소(본점), 사업장(지점), 개업년월일, 사업의 종류, 소득금액 등

<부록 3> 4대 사회보험 주요 수록정보

구분	수록 정보
4대 사회보험 공통	사업장 기호, 사업장 명칭, 형태, 소재지, 전화번호, FAX번호, E-mail, 휴대전화, 업태, 종목, 주생산품, 업종코드, 사업자등록번호, 법인등록번호, 사용자(대표자) 내역(성명, 주민등록번호, 주소, 전화번호) 등
국민연금	사업자 사항: 근로자수, 가입대상자수, 분리적용사업장 해당여부, 본점내역 등 지역가입자 사항: 성명, 주민등록번호, 세대주와의 관계, 주소, 전화번호, E-mail, 월소득액, 자격취득연월일, 업종 및 직업명, 업종코드, 자영자 영업형태(종업원 수, 경영면적, 경영구분, 영업장 소재지) 등
건강보험	사업자 사항: 단위사업장수, 적용대상자수, 사업장 특성부호, 적용연월일, 봉급지급일 및 회계 등 지역가입자 사항: 세대주(성명, 주민번호, 주소, 전화번호 등) 및 가입자 내역(관계, 성명, 주민등록번호, 취득일자, 장애인/국가유공자 부호/등급 및 등록일) 등
고용보험	상시근로자수, 실업급여 피보험자수, 적용사업(전사업, 실업급여), 성립일, 사무조합명칭, 사무조합번호, 주된 사업장의 명칭, 소재지, 업종, 총상시 근로자수, 총사업장수, 사업자등록번호, 대규모기업, 실업급여피보험자총수, 적용사업, 법인등록번호, 주된 사업장관리번호, 전화번호, 주생산품, 업종코드 등
산재보험	상시근로자수, 사업의 종류, 사업종류코드, 사업의 형태, 사업기간, 성립일, 사무조합명칭, 사무조합번호 등

<부록 4> 기타 행정자료 주요 수록정보

구분	수록 정보
운전면허 (경찰청)	등록번호, 성명, 주민등록번호, 주소, 적성검사기간, 면허발급일 등
여권 (지자체)	성명, 영문 성명, 성별, 주민등록번호, 생년월일, 주소, 여권번호 등
국가자격증 (산업인력공단)	성명, 주민등록번호, 주소, 자격 내역 등
외국인등록 (법무부)	외국인 등록번호, 출신국가, 비자 종류, 한국에서의 주소 등