

미국 센서스의 변화와 향후 전망: 2000년의 경험과 2010년의 계획을 중심으로*

전 광 희**

이 글은 1990년부터 진행된 미국의 인구센서스의 변화와 향후의 전망에 대해서 2000년의 경험과 2010년의 계획을 중심으로 논의한다. 먼저, 2000년 센서스의 준비과정에서 미국 과학원학술평의회 국가통계위원회가 작성한 2000년 센서스 개선안의 내용을 통계적 추정방법 도입, 응답률 개선방안, long form 개선방안, 소수민족 집계누락, 인구센서스 대체방안, 센서스 비실시 연도 소지역통계 확보방안, 센서스 프레임 정비 등으로 살펴보고, 1998년 대법원 판결 이후 2000년 센서스에서 통계적 추정방법을 이용한 미국판 One Number Census를 추진하지 못하게 된 경위를 검토한다. 다음은 2010년 인구센서스의 핵심적 특성의 하나인 long form과 short form의 분리, 곧 short form은 전통적 센서스, 곧 전수조사와 현장실사의 원칙을 준수하는 방식으로 그대로 두고, long form을 대체하여 순환센서스로 진행되는 American Community Survey에 대해서는 그 과정과 주요 특성을 검토한다. MAF/TIGER는 인구센서스의 집계과정에서 정확성을 향상하고자 하는 또 하나의 특성으로, 이것은 2010년 센서스에서 모바일 컴퓨터를 도입하는 하이테크 센서스를 목표로 기획된 것이었다. 이 글은 미국 센서스에서 short form과 long form의 분리가 대법원 판결로 우연히 일어난 것이 아니라는 점을 강조하기는 하지만, 미국 센서스는 2008년 현재, 인구조사국의 모바일컴퓨터 개발실패로 60만 비정규직 조사원의 고용을 불가피하게 할 것으로 예상되며, 이로 말미암아 2010년 미국 센서스의 정확성에 대한 우려감이 증대하고 있음을 지적한다.

핵심 단어: 미국 센서스, 인구조사국, short form, long form, MAF/TIGER, American Community Survey, One Number Census, 통계적 추정방법, 모바일 컴퓨터, 순환센서스, 표본센서스, 센서스 대체방안

* 본 연구는 저자가 2007-2008년 충남대학교의 재정지원으로 미국 University of Wisconsin-Madison 방문교수로 파견 중 수행된 연구의 일부로, 초고는 한국사회학회 인구도시연구회 2008년 동계세미나에서 발표되었으며, 세미나 참석자들의 의견을 일부 반영하였다.

** 충남대학교 사회학과 교수 | E-mail: jkh96@cnu.ac.kr

I. 머리말

미국은 1787년에 채택된 연방헌법의 규정에 의거하여, 하원의석배정, 연방자금배정 등을 주목적으로 인구센서스를 실시하고 있다. 그러나 미국의 인구센서스는 20세기가 시작되기 전에는 총인구를 단순 집계하는 수준에 불과하였다. 그 후 미국 연방법을 재13편(US Code Title 13)에 의거하여, 1903년에 인구조사국(Census Bureau)이 설치되었으며, 인구조사국은 1910년부터 10년을 주기로 하는 이른바 "Decennial Census"를 주도하는 핵심적 국가기관이 되었다. 인구조사국은 1940년 이후 실시된 인구센서스에서 과소집계 수준을 추정하는데 관심을 보이기 시작하였다(Anderson, 1988: 221). 또 인구조사국은 1990년에는 종전의 인구분석(demographic analysis)을 통한 센서스 인구의 정확성 평가(accuracy assessment)와 더불어, 사후조사(post-enumeration survey)라 불리는 대규모의 표본조사를 실시하여 센서스 인구의 과소집계 구조를 체계적으로 파악하기 시작하였다(Edmonston and Schultze, 1996: 30-31).

1990년 인구센서스의 막대한 조사비용과 과소집계¹⁾, 특히 소수민족의 심각한 집계누락은 정치적 쟁점으로 비화하였으며, 1992년에는 연방의회가 마침내 미국과학원학술평의회(National Academy of Science National Research Council)의 위탁사업으로 인구센서스를 근본적으로 재검토할 것을 지시하는 법률을 통과시켰다(Edmonston and Schultze, 1996: xiii). 연방의회를 통과한 PL 102-125 '1991년 센서스 개선법'(Decennial Census Improvement Act of 1991)에 의거하여, 학술평의회는 국가통계위원회(Committee on National Statistics)는 2000년 이후에 실시되는 인구센서스에서 (1) 인구수를 가능한 정확하게 집계하는 방법 (2) 이와 일치하는 인구·주택 관련 자료수집의 대체수단에 대하여 검토할 것을 요청하였다(Edmonston and Schultze, 1996: xiii)

PL 102-125는 학술평의회가 설치한 패널에서, 미국시민의 사생활 보호의식, 인구센서스에 대한 신뢰도, 정치적 변수 등의 요인이 센서스의 중립성에 주는 영향에 관한 분석을 포함하여, 전수조사(全數調査)에 입각한 전통적 인구센서스 방식에 대한 다양한 대체적 수단의 장단점과 비용경감 효과를 비교검토하고 이것을 토대로 하여 권고안을 채택해 줄 것을 요구하였다.

1) 1990년 센서스는 종전의 센서스보다 많은 경비를 지출하였지만, 집계누락으로 인한 센서스의 정확성 문제는 1980년 센서스보다 훨씬 더 심각했던 것으로 보고되고 있다(Edmonston and Schultze, 1996: iii).

검토과제를 구체적으로 보면, (1) 조사대상자에 대한 직접적 자료수집방법의 개선책을 포함하는 집계방법의 전반적 개선책 (2) 행정등록자료, 지역조사, 표본조사 등에서 얻을 수 있는 정보의 축적 또는 순환방식(rolling method)에 의한 자료수집을 포함하는 인구집계에 필요한 대체적 방법의 가능성에 대한 검토, (3) 주, 지구, 센서스 조사구, 센서스 블록 등의 다양한 지역수준에서 수집한 자료의 정확성에 대한 검토를 포함하여, 인구자료의 수집과 함께 정확성을 개선하기 위하여, 표본조사방법을 센서스 본조사의 현지실사 방식과 연동하는 것에 대한 타당성 검토 (4) 인구집계 결과와 함께, 향후 여타 인구센서스 조사 항목 자료에 대한 수요의 변화에 대한 전망적 검토, (5) 예상가능한 자료의 수요변화에 대하여, 종래의 조사방식보다 상대적으로 우위에 있는 자료수집방법의 가능성 및 대체적 정보원과 그것의 입수방법 검토 등의 많은 검토과제를 포함하였다. 또 해당 패널에는 이러한 개별 검토과제에 대하여, 2000년 인구센서스 실시시점까지 신뢰할 수 있고 유효한 방법으로 확립될 수 있는 것이 무엇인가를 판단하는 것도 과제대상에 포함되어 있었다(Edmonston and Schultze, 1996: xiii-xiv).

PL 102-125가 연방의회를 통과하면서, 미국 학술평의회는 국가통계위원회 산하에 '2000년 이후 센서스 문제에 관한 패널'(Panel on Census Requirements in the Year 2000 and Beyond)(간단히, '패널'로 부름)이 설치하였다.

II. 패널보고서와 미국 인구조사국의 대응

1. 2000년 이후 인구센서스에 대한 패널의 보고서 내용

패널보고서는 2000년 이후 미국 센서스의 미래를 장기적으로 전망하면서 몇 가지 과제를 제안하고 있다. 보고서의 권고안은 아래와 같은 위원회 심의의 결론을 바탕으로 하고 있다(Edmonston and Schultze, 1996: 3).

(1) 전수조사에 의한 전통적 집계방법을 토대로 하여, 최후의 1인까지 집계하는 센서스 방식은 실효성에 의문이 제기되며, 단순한 예산증액조치로 종래형의 인구센서스를 실시한다고 하더라도, 집계의 정확성 또는 자료의 품질이 괄목할만한 수준으로 개선된다고 보기 힘들다.

(2) 전통적 센서스의 전수조사를 통한 집계방법을 축소하여, 통계적으로 집계누락의 규모나 특성을 추정함으로써 대부분의 주요 변수에 대한 집계의 정확성을 개선할 수 있다.

(3) 인구수의 집계에서 통계적 추정을 사용하기로 결정하는 경우, 센서스의 실시방법과 운영절차를 재설계함으로써, 2000년 이후 인구센서스의 소요경비를 삭감할 수 있으며, 나아가 집계의 정확성을 개선하는데도 기여할 수 있다.

(4) 2000년 이후 센서스부터는 long form을 폐지하고, 해당 자료의 확보수단으로 매월 실시되는 표본조사(나중에, 이 조사는 American Community Survey라고 명명됨)를 이용하려는 제안은 그것의 영향과 예상경비를 정확하게 평가하기 위하여 새로운 연구와 작업이 필요하다. 소지역 통계자료를 더욱 빈번하게 확보할 수 있는 연속적 관측법(continuous measurement)에 대해서는 진지하게 검토하여 볼 필요가 있다. 그러나 필요한 연구와 평가의 작업은 2000년 센서스의 실시시점까지 완료될 수 없을 것으로 예상되기 때문에, 2000년 센서스에서 종전대로 long form 조사표를 조사계획에 포함시킬 필요가 있다.

아래는 패널의 보고서가 미국의 2000년 이후 센서스를 재설계하기 위하여 제안한 권고안 중에서 본 연구의 전개에 필요한 부분을 좀 더 상세하게 정리한 것이다.

1) 통계적 추정방법의 도입

조사원 방식에 의한 전통적 인구집계의 실정은 인구센서스가 추구하는 집계오차의 허용한도를 넘어서고, 특히 소수민족의 집계누락은 심각한 수준에 이르고 있다. 이 문제에 대해서, 패널은 우편조사와 추계방식을 조합함으로써, 비용의 삭감과 정확성의 개선이라는 효과를 달성할 수 있다는 결론에 도달하였다(Edmonston and Schultze, 1996: 3-4).

추정방식의 도입에 대해서는 패널은 Recommendation 5-1에서, 적절한 노력 후에 일정시점에서 조사 불응자의 직접집계를 중단하고, 표본추출에 의하여 남아있는 조사 불응자와 그들의 특성을 추정할 것을 주문하고 있다. 또 Recommendation 5-2에서는 지리적 지역 또는 인구특성 집단 간에 발생하는 과소집계 수준의 차이를 축소하기 위하여, 직접적 집계와 조사 불응자에 대한 표본조사를 토대로 하여 추정결과를 개선하고, 가용자원을 최대한 이용하여 센서스에서 집계결과를 최고의 단일수치(the best single-number)로 조정할 것

을 인구조사국에 주문하고 있다.

2) 응답률의 개선방안

패널보고서는 응답률의 개선방안으로, 먼저 Recommendation 5-3에서, 인구조사국에 조사 대상자가 이해하기 쉬운 조사표의 채택과 연방 법률에 규정된 센서스에 대한 응답의무를 광범위하게 홍보할 필요성을 제안하고 있다. Recommendation 5-4에서는 센서스 조사표를 우편으로 전달하기 위한 주소파일(MAF: Master Address File)을 개선하기 위하여, 인구조사국에 대해서는 주정부와 여타 지방정부의 제휴를 강화할 것을 주문하고, 또 미국의회에 대하여 우편파일을 통계목적으로, 주정부, 지방정부, 연방통계기관 사이에 공동사용이 가능하도록 미국 연방법률 제13편(Title 13)을 개정할 것을 주문하고 있다. 또 패널보고서는 미국의 우정사업청(Postal Service)과 인구조사국이 상호협력하여 센서스의 개선을 위하여 노력할 것을 권고하고, 특히 우정사업청에 대해서는 거주주택의 주소파일을 갱신·개선하도록 하고, 또 우편조사를 통한 직접 집계에서 정확성을 개선하기 위하여 센서스 실시 전에 주소 리스트의 점검·우송, 센서스 실시기간 중에는 거주주택이 비어있는가 또는 아닌가를 확인하는 작업을 실시할 것 등을 주문하고 있다(Edmonston and Schultze, 1996).

3) Long form의 개선방안

패널보고서는 1990년 인구센서스에서 long form이 응답률의 저하와 예산증액의 원인이 되고 있다는 인식을 강화하여, Recommendation 6-1에서 long form 대신에 조사내용을 몇 종류의 조사표로 분할한 중간적 조사표(intermediate forms)를 이용하는 매트릭스 표집방식(matrix sampling method)의 도입에 대하여 다양한 평가를 실시할 것을 주문하고 있다(Edmonston and Schultze, 1996: 8-9).

한편, 패널보고서는 long form과는 별도로 인구조사국이 독자적으로 도입을 검토하고 있었던 월별 대규모 표본조사에 대하여 큰 관심을 보이고 있지만, 인구조사국이 이 조사의 경비절감 효과를 과대평가하고 있으며, 또 추정방식이나 소지역별로 누적된 자료의 사용에 대하여 향후 검토해야 할 과제가 많이 남아있다는 점에 주목하였다(Edmonston and Schultze, 1996). 그 결과, 패널보고서는 2000년 인구센서스에서는 long form을 폐지하고, 이를 대체하는 다른 표본조사를 채택하

는 안을 권고하지 않기로 하는 결론에 도달하였다(Edmonston, 1995: 9-10). 이를 받아들여, 패널보고서는 Recommendation 6-2에서 연속형 표본조사를 포함하는 몇 가지 대체적 조사방법의 경비절감 효과를 비교검토하고, 이들 조사방법이 전국적 가구조사 시스템에 통합적으로 운용될 수 있는 가능성을 지속적으로 연구·검토할 것을 주문하고 있다(Edmonston and Schultze, 1996: 10).

이상에서 본 것처럼, 패널은 2000년 인구센서스에서는 종전대로 long form에 의한 전수조사를 실시하고, 인구조사국이 도입을 계획하고 있는 월별로 진행되는 대규모 표본조사에 대해서는 그 효율성을 지속적으로 점검할 것을 주문하고 있다.

4) 소수민족의 집계누락

미국은 다민족 국가로서, 1960년대부터 인종차별과 소수민족의 문제에 대한 미국 시민의 관심이 점증하는 가운데, 미국 내에서 인종·민족별 인구구성에 대한 정확한 자료의 수집은 사회적으로 그 수요가 증대하였다. 1965년의 선거권법(Voting Rights Act of 1965)과 법정통역(court interpretation) 제도의 도입과 더불어, 하원 선거구 획정과 각종 보조금 프로그램의 입안·집행을 위하여, 소지역 수준의 인종·민족별 인구구성에 대한 자료에 대한 수요가 급격히 증가하였다. 또, 1990년의 인구센서스에서는 각 인종·민족별 출신 간에 과소집계의 차이가 훨씬 증가하여 심각한 수준에 이르렀다. 이러한 사정들은 2000년 이후, 인구센서스에서 소지역 수준에서 인종·민족별 인구구성에 관한 자료 확보의 필요성이 증대하는 중대한 요인이 되기도 하였다.

1990년의 인구센서스에서 인종·민족별 출신에 대해서, 15개의 중분류와 소분류가, 히스패닉 계통에 대해서는 4개의 중분류가 채택이 되었다. 인구센서스에서는 응답자 자신이 “주관적으로, 자기의 인종 또는 민족별 출신을 어떻게 지각하는가?”라는 문항에 대하여 응답자가 선택할 수 있도록 하는 구조화된 조사표를 이용하였다. 국경을 넘어 유입되는 이민이 지속적으로 증가하는 가운데, 미국의 개별 가구가 인종·민족별 구성에서 한층 다양화하고, 인종·민족 간에 혼인도 증가하는 가운데, 다수의 응답자들이 점차적으로 복수의 인종적 자각을 가지게 되는 것이 일상화되기 시작하였다.

인종·민족별 구성이 근년에 이르러 훨씬 더 유동성을 증대하고 있는 사실을 근거로, 패널보고서는 이 조사항목이 본래 다의적(多義的)이라는 인식을 근거로 하여, 그것을 반영하는 것이 정치적으로나 사회적으로나 행정적으로도 유효

한 자료를 획득하는 조사방법에 대하여 신중한 검증방법이 채택될 필요가 있다고 주장하고 있다(Edmonston and Schultze, 1996: 11)

이러한 주장을 받아들여 패널보고서는 Recommendation 7-1에서, 인구조사국에 인종·민족별 출신의 정의가 응답자에게 쉽게 받아들여질 수 있고, 인구센서스의 여타 항목과 관계가 명확하여 제표화하는데 효과적이고, 센서스 자료와 여타 연방정부통계 또는 행정등록자료에서 이들 항목간의 비교가 가능하고, 또 지역과 특정의 인구집단에 관한 양질의 자료가 확보될 수 있도록 하기 위하여 질문방법에 대하여 다양한 주문을 하고 있다(Edmonston and Schultze, 1996: 11-12). 연방행정예산청(OMB: Office of Management and Budget)이 작성하는 통계작성지침 15호(Statistical Directive 15)가 인구센서스 자료를 근거로 재검토의 대상이 되면서, 인구센서스의 인종·민족별 출신에 대해서 정확성이 높은 자료가 특별히 요청되기에 이르렀다.

이상에서 본 것처럼, 패널보고서는 Recommendation 7-2에서, 연방행정예산청에 대하여, 인구조사국이 2000년 인구센서스의 인종·민족별 출신에 대한 조사방법을 결정할 때에, 사전에 충분한 검토기회를 가지도록 하고, 시간적 여유를 가지고 통계작성지침 15호를 공표할 것을 요청하고 있다(Edmonston and Schultze, 1996: 12).

5) 인구센서스의 대체방안

패널보고서는 2000년 이후 장래의 인구센서스에 관한 검토결과를 발표하였는데, 그 결과는 다양하고 포괄적인 내용을 담고 있었다. 전통적인 집계방법에 의한 센서스 본조사와 표본조사에 의한 추정결과를 종합하는 인구센서스 계획을 최종적으로 권고하였지만, Recommendation에 이르는 과정에서 몇 가지 다른 대안들도 고려의 대상이 되었다. 또 패널은 인구센서스의 대체방안을 검토하면서 센서스 자료가 충족해야 요건으로서 다음 3가지를 설정하였다. 곧 (1) 연방 헌법, 연방법을 그리고 여타 법적 이유로 가구와 개인에 대한 특성별 정보는 지리적 위치, 주소지 정보와 연계되어야 하며 (2) 특히 투표권법의 시행 이후 주소지와 연계하여 인종·민족별 출신을 분명히 할 수 있어야 하며, (3) 광범위한 인구특성정보를 가진 소지역에 대한 통계자료를 제공할 수 있어야 한다(Edmonston and Schultze, 1996: 12)

이상과 같은 전제를 바탕으로, 패널보고서는 주민등록제도(national register system) 또는 행정등록자료(administrative records)를 토대로 하는 센서스, 우

정사업청을 실시기관으로 하는 센서스, 표본센서스(sample census) 또는 순환 센서스(rolling census) 등의 대체방식을 다양하게 검토하였다. 이와 관련된 검토내용과 결론을 요약하면 다음과 같다.

먼저, 패널보고서는 주민등록에 대해서 다음과 같은 이유로 인구센서스의 대안이 될 수 없다고 주장한다. 먼저, 미국은 현재 주민등록제도를 채택하지 않고 있으며, 그것을 도입하자는 제안도 사생활 보호나 시민권의 관점에서 민감한 문제로서, 사회문제로 비화될 것이 분명하며, 미국의 문화적 관점에서도 받아들여지기 힘든 제도로 취급하고 있다. 주민등록을 인구센서스의 대체수단으로 사용하자는 제안에 대해서는, 미국 국민의 흔쾌한 협력을 기대하기 힘들며 결과적으로 좋은 품질의 자료를 확보하는 것이 사실상 불가능하다는 것이 그 이유이다.

행정등록자료에 의한 센서스의 가능성에 대한 패널보고서의 검토결과는 다음과 같은 것이 있었다. 센서스에서 집계가 곤란한 인구특성 집단의 소재정보원으로서, 부분적으로 행정정보가 사용되고 있다. 그러나 행정등록자료로 센서스를 전면적으로 대체하는 방법에 대해서, 패널보고서는 (1) 기존의 행정정보가 어떤 경우에도 단독으로 센서스의 전체 조사항목을 포괄하지 못하고, (2) 이들의 작성과정에서 행정기록 상호간에 대규모의 링크(link)가 필요하며, (3) 링크에 필요한 소요예산의 증가나 조사대상자의 사생활 보호라는 면에서 문제가 생기고, (4) 행정등록자료가 보유하고 있는 단편적인 인종·민족별 출신에 관한 자료에서는 신뢰성의 문제가 생기고, (5) 통일된 포괄적 식별정보(예: 주민등록번호)가 없는 상태에서 행정등록자료를 링크할 때 중복배제의 문제를 해결하기 힘들다고 하는 다양한 문제를 지적하고 있다. 그 결과, 패널보고서는 2000년 인구센서스를 복유립에서 현재 실시되고 있는 행정등록자료를 바탕으로 하는, 이른바 "등록센서스"(register-based census)로 실시하는 것은 현실적으로 불가능하다는 결론을 내렸다(Edmonston and Schultze, 1996: 13).

다음은 센서스의 기획과 통계처리는 인구조사국이 하고, 조사표의 배포와 반송에 관련된 업무는 우정사업청의 우편집배원에게 위탁하는 시스템의 실현가능성에 대하여 인구조사국과 우정사업청 사이에 연구되었다. 그러나 현지실사의 추적(follow-up) 과정에서 표본조사를 통하여 보정이 요구되는 조사 불응자에 관한 센서스 정보를 어떻게 확보할 것인가에 관한 문제는 미해결의 상태로 남아 있었다. 우정사업청의 우편집배원의 급료가 인구센서스 조사원의 3배 이상이라는 것도 문제로 지적되었다. 패널보고서의 결론은 우편집배원에 의한 현지실사의 실시방안은 현실성이 없다는 것이었다.

다음으로, 패널보고서는 "표본센서스"와 "순환센서스"가 전통적 센서스의 대

체방안이 될 수 있는가에 대해서도 검토하고 있다.

표본센서스는 전통적 센서스 대신에 특정시점에서 대규모 표본조사를 실시하고, 그것을 근거로 인구의 지역별, 특성별 규모, 그리고 다양한 특성별 정보를 추정치로 획득하는 것이다. 표본센서스에 관련된 최대의 문제는 미국 연방대법원이 센서스의 기본요건으로서 특정시점에서 총인구를 현장 실사할 것을 요구한다는 것이었다. 또, 표본조사는 전수조사에 비하여 조사 불응자의 비율이 상대적으로 높을 뿐만 아니라, 그 비율이 조상대상자의 인구학적 특성별로 균일하지 않다는 특징이 있다.

한편, 패널보고서는 매년 모집단의 10%를 표본으로 추출하여 조사하는 소위 “순환센서스”에 대한 검토의견을 제시하였다. 패널보고서의 순환센서스에 대한 검토결과는 다음과 같았다. 곧, 순환센서스 방식은 센서스 소요예산을 연차별로 평준화한다는 점에서는 장점이 있지만, 예산의 누적액수로 판단한다면, 전수조사와 현장실사를 대원칙으로 하는 종전의 전통적 센서스보다 오히려 실시비용이 많이 든다는 사실을 간과하지 말 것을 주의하고 있다.

유럽의 일부 국가에 실시되고 있는 표본센서스나 순환센서스의 어느 것이든 종래의 전통적 센서스 방식과 비교할 때, 그것이 센서스의 원칙을 위반하는 것은 사실이며 나아가 센서스 방법론의 일대혁신이라는 것이 분명하기는 하지만, 이와 함께 국민의 센서스에 대한 협력수준의 저하는 불가피할 것이라고 전망하고 있다. 이러한 이유 때문에, 패널은 이들 조사방식 중에서 어떠한 것도 2000년 인구센서스에서 채택될 수 있는 대체방식으로 받아들이기 어렵다는 결론을 내리고 있다(Edmonston, 1996: 14).

6) 센서스 비(非) 실시연도의 소지역 통계자료의 확보방안

미국 인구센서스의 경우 자료이용이라는 점에서 가장 큰 특징 중의 하나는 그 자료가 연방정부나 주정부만이 아니라, 시, 군, 학구 등의 소지역 수준에서 통계자료를 제공할 수 있다는 점이다. 패널보고서는 미국 인구센서스의 미래에 대하여 검토를 시작하던 당시, 이미 인구조사국에서는 미국 전역을 대상으로 매년 대규모 표본조사를 실시하고, 인구규모나 상대적으로 적은 지역이나 인구 집단에 대해서는 3년 또는 5년간의 자료를 누적하여 그것의 이동평균(moving average)을 구하여 사용함으로써, long form 대신에 매년 소지역 수준의 통계자료를 제공한다고 하는 소위 “센서스 long form의 재편계획”을 구체화하고 있었다. 이러한 계획에 대하여, 패널보고서는 현시점에서는 소지역 통계자료의 이용

자료부터 충분한 동의를 얻지 못하고 있기 때문에, 인구조사국은 통계 이용자의 반응을 신중하게 검토할 필요가 있다는 결론에 도달하였다. 이러한 관점에서, 인구조사국이 구체화하고 있는 long form을 대체하는 새로운 조사를 포함하여, 센서스의 비 실시연도에 대하여 소지역 통계자료의 품질, 양, 빈도 등의 개선을 인구조사국에 요청하고 있다(Edmonston and Schultze, 1996: 15).

7) 센서스 프레임의 정비

미국 인구조사국은 우편조사방식을 도입한 이후부터, 센서스 실시에 즈음하여 거주가구의 주소 리스트, 곧 Master Address File (MAF)를 작성하여 왔다. 패널이 설치되던 당시, 인구조사국은 연속적으로 업데이트되는 MAF를 사용하여 센서스자료, 행정등록자료, 기타 통계자료 등에서 얻은 정보를 바탕으로, 센서스 비 실시연도에 대해서 소지역 자료의 정비를 시도하는 전략을 수립함과 동시에, 우정사업청의 협력으로 본격적으로 MAF의 상시 데이터베이스 구축사업에 착수하였다.

패널은 이러한 MAF 상시 데이터베이스 구축사업이 진행 중이라는 보고를 접수하고, Recommendation 8-3에서 이러한 사업이 인구센서스 실시 이전에 주소 리스트를 재정비하는 종전의 방식과 비교할 때, 경비절감효과가 기대되는 경우는 물론 종전의 방식보다 많은 경비가 소요되는 경우에도 MAF의 정비가 가져다주는 파급효과가 분명한 경우에는 개발사업을 지속적으로 추진하여야 한다고 주장하여, 상시 데이터베이스 구축사업을 적극적으로 평가하였다. 이와 함께, 패널은 같은 Recommendation 8-3에서 MAF의 안전관리를 포함하여 통계작성이 목적인 경우 연방정부의 통계기관은 물론 주정부나 여타 지방정부도 사용이 가능하도록 하는 추가적 조치가 필요하다고 주장하고 있다.

2. 2000년 인구센서스 실시를 위한 미국 인구조사국의 대응

미국 인구조사국은 과학원학술평의회 국가통계위원회의 이러한 패널보고서를 접수하고, 2000년 인구센서스의 본조사를 실시하고, 이른바 “통합적 범위측정조사”(Integrated Coverage Measurement Survey, 약칭으로 보통 ICM을 사용함)를 대규모 표본조사로 실시하여, 본조사에서의 과소집계의 정도를 지리적 지역·인구학적 특성별로 평가한 이후에, 이것을 이용하여 본조사의 결과를 보정한 것을 정규의 인구센서스 최종적 결과수치로 확정하는 인구센서스 실시사

업을 추진하였다. 물론, 본조사의 결과수치를 대규모 표본조사나 기존의 행정기록 등으로 보정함으로써, 정확성이 높은 최종적 결과수치를 생산하는 센서스 방법론은 미국에서만 생겨난 독창적 아이디어는 아니라고 할 수 있다. 영국에서도 1991년 인구센서스의 실패를 반성하면서, 본조사에서 얻은 결과수치를, 본조사 후에 과소집계의 정도를 평가하기 위하여 실시한 “센서스 범위조사”(CCS, Census Coverage Survey)로 명명된 대규모 표본조사의 결과와 각종의 행정등록자료의 정보를 통합함으로써, 인구센서스로서 단수(單數)의 최종적 결과수치 세트를 생산하는 소위 “One Number Census”(ONC)에 의한 인구센서스 개선방안을 수립하고, 그것을 구체화하기 이르렀다. 영국은 당초의 계획대로 2001년 인구센서스는 ONC 프로젝트의 틀 안에서 진행하였으며, 센서스 범위조사와 각종 행정등록정보에 의하여 보정·추정된 수치로서 “one number”(단수, 유일의 수치)가 인구센서스의 각종 제표작업에서 최종적 결과수치로 사용되고 있다.

그러나 미국의 인구조사국이 2000년 인구센서스의 실시에 즈음하여 추진하였던 미국판 ONC는 예기치 않은 좌절을 맛보게 되었다. 연방대법원은 1998년 1월 25일, 표본조사를 토대로 통계적 추정방법을 이용하여 본조사의 결과를 보정한 수치를 연방하원의 의석배분에 사용하는 근거숫자로 사용하는 것은 연방법률을 위반한다는 최종판결을 내렸다²⁾. 이러한 연방대법원의 판결은 인구조사국에 대하여, 범위추정조사의 결과활용과 관련하여 기존계획의 근본적 방향수정을 요구하기에 이르렀다. 그 결과, 인구조사국은 2000년 인구센서스를 패널보고서의 Recommendation과는 거리가 먼 전통적 조사방식에 따라, 전수조사와 현장실사의 원칙을 고수하기로 하는 결정을 내렸다.

Ⅲ. 2000년 센서스의 문제점과 2010년 센서스의 과제설정

2000년 인구센서스는 1940년 이후 실시된 인구센서스들 중에서 집계적 정확성이라는 면에 품질이 가장 우수한 조사결과를 얻었다는 평가를 받고 있다. 그

2) 연방대법원의 판결 결과, 인구조사국은 선거구의 구획수정과 하원의원 정수의 재배정을 목적으로 센서스 본조사의 집계결과가 아닌 다른 수치, 결국 통계적 추정방법(statistical estimation technique)을 이용하는 수치를 제공할 필요가 없었다. 그러나 인구조사국은 연방대법원의 판결이 인구추계방법의 개선 등의 조사연구에 이르기까지 센서스 집계결과와 보정수치를 사용해서는 안 된다는 판단을 하지는 않았다고 할 수 있다.

러나 2000년 인구센서스가 아무런 문제가 없었다고 생각해서는 안 될 것이다³⁾. 고액의 조사경비, 인종·민족별 출신에 관한 자료획득의 어려움⁴⁾, 유동인구·시설인구·무주택(가출, homeless)인구 집계 문제점, 인구센서스 결과(특히, long form)의 지연공표로 인한 문제점, 불완전한 조사구 지도정보, 응답종료 가구에 대한 조사원들에 의한 조사의 중복실시, 집계용 기계의 문제 등 수많은 과제가 2000년 인구센서스의 새로운 쟁점으로 부각되었다.

이러한 2000년 인구센서스 실시에 관한 문제점의 인식을 바탕으로, 집계누락과 중복집계의 배제를 통한 정확성의 개선, 조사결과의 신속한 제공, 작업위험의 삭감, 센서스 재검토에 의한 경비절감이라는 한 세트의 과제들이 다음 인구센서스를 목표로 하여 설정되었다. 특히, 2010년 인구센서스 실시를 앞두고, 추가적으로 검토대상이 된 것은 다음과 같은 것들이 있다. ① 현지실사에서 모바일 컴퓨터의 사용으로 문서작업업무와 이동거리의 삭감, 조사 불응자에 대한 독촉업무의 효율화, GPS에 의한 지역안내정보의 활용, 중복조사의 회피 ② 인종·민족에 관한 조사항목의 개선, 조사표 문항의 표현방식을 연구하여 히스패닉 등 소수민족의 집계 정확성 향상 ③ 집계 정확성의 향상을 위하여 현재인구방식의 검토 ④ 웹방식 또는 전화방식의 채택을 염두에 둔 현지실사 방법의 개선 ⑤ 스페인어권 지역에서 영어·스페인어 조사표의 테스트 ⑥ 센서스 조사구 내의 일반거주지와 집합거주시설의 구분 ⑦ 이름, 생년월일 대조를 통한 학생 등의 중복기재 방지책의 검토가 그것이다.

IV. 2010년 센서스 설계의 핵심요소

미국에서 2010년 실시하기로 예정되어 있는 센서스의 조사계획에는 두 개의 핵심요소가 있는데, 바로 ① 센서스 본조사의 short form과 long form의 분리

3) 우편응답률은 전국 수준에서 mailout/mailback 방식이 처음으로 도입되었던 1970년 센서스의 경우가 78%에 이르렀다. 그러나 1980년 센서스에서는 75%, 1990년에는 65%로 응답률이 지속적으로 감소하였다. 2000년 센서스에서는 short form과 long form을 합쳐서 우편응답률이 74%로 1990년 센서스보다는 상당정도 개선되었다.

4) 연방행정예산청은 인종구성의 변화를 토대로 정기적으로 인종의 분류기준을 재검토한다. 이 기관은 2000년 센서스 직전에 분류기준을 개정하였다. 이것을 토대로 2000년 센서스에서는 조사대상자들이 5개의 인종구분 중에서 하나 이상의 인종 또는 "기타"를 선택하는 형식으로 조사가 진행되었다. 그 결과, "기타"를 선택하는 응답자가 대폭적으로 증가하였으며, 인구조사국은 조사문항의 언어재구성을 통하여 다양한 가능성을 시험조사에서 검증하고 있다(Census Bureau, 2004: 7)

② 조사프레임인 가구명부의 정비와 조사구 지도정보의 상시 데이터베이스화가 그것이다.

1. 센서스 본조사의 short form과 long form의 분리

미국의 인구센서스에서 long form은 1940년에 처음으로 도입되었으며, 인구학적 특성, 주택, 사회경제적 특성에 관한 자료를 수집할 목적으로, 일부의 가구를 조사대상으로 하는 표본조사의 형태를 취하였다. 인구조사국은 1940년부터 2000년까지 7회에 걸쳐, short form과 long form이라는 두 종류의 조사표를 이용하여 조사를 실시하였다. 예컨대, 2000년 센서스의 short form (Form D-61A)의 조사사항은 ① 성씨 ② 이름 ③ 성별 ④ 연령 및 생년월일 ⑤ 히스패닉·라틴계 여부 ⑥ 인종 또는 민족별 출신 등 6개 항목으로서, 한국의 2005년 인구주택총조사의 전수조사 항목과 비교할 때 조사항목수가 비교가 안 될 정도로 적은 편이다. 이에 반하여, long form (Form D-61B)은 short form의 6개 조사항목을 포함하며, 최초의 가구주 1명에게는 가구관련 문항을 포함한 53개의 조사항목, 가구원에 대해서도 32개 조사항목으로, 6명의 가구원까지 연속 기재가 가능하도록 만들어져서, 분량이 총 37쪽이나 되는 꽤나 두터운 조사표라고 할 수 있다. 또 1990년과 2000년에는 전국 평균 6개 가구 중 1개 가구(17%의 표집률)가 long form 조사의 대상가구였지만, 인구규모가 2,500명 이하인 지역에서는 2가구 중 1가구가 long form 조사의 대상가구가 되었다.

Long form은 조사항목이 많은 것만이 문제가 아니라, 가구주 관련항목 중에는 공공요금지출액, 가구소득, 주택평가액과 같이 정확하게 응답하기 위해서는 과거의 금전출납부를 재점검해야 할 필요성이 있는 것이 있었다. 개인정보에 관해서도, 학력, 신체적 조건, 수입의 종류와 금액 등 많은 문항들이 조사대상자들이 응답을 기재할 때, 조사기피를 야기할 수 있는 항목들이 상당수 포함되어 있었다⁵⁾. 이 때문에, long form 조사표는 short form 조사표와 비교할 때 응답률이 낮은 것이 특징이며, 이것은 인구센서스의 정확성을 떨어뜨리는 요인의 하나가 되어 왔다.

미국 인구조사국의 2010년 센서스 계획에 의하면, 1940년 센서스부터 실시되어 왔던 long form을 센서스에서 분리하여, 센서스를 short form으로만 실시하기로 하였으며, 그 준비는 2008년 현재 거의 완료된 상태라고 할 수 있다. 2010

5) 클린턴 행정부에서 인구조사국장이었던 Kenneth Prewitt는 2000년 4월 11일 기자회견 발표문에서, long form의 조사 불응자 비율이 1990년 센서스에서의 short form의 조사 불응자 비율의 2배 이상의 규모에 이를 것이라고 예상하였다.

년 센서스에서 long form을 분리하는데 대해서, 우선 센서스에서 상대적으로 기피경향이 강한 long form을 센서스 본조사에서 완전히 분리하여, 센서스를 전적으로 인구수의 파악이라는 연방헌법의 요구사항을 준수하는 것으로 특화된 조사로 실시하여, 조사대상자의 협력수준을 개선하는데 초점이 맞추어 지고 있다. 이것은 어떤 의미에서 미국의 인구센서스를 1940년 센서스 이전의 상태로 복원하는 것이라고 할 수 있을 것이다.

미국의 2010년 센서스에서 short form과 long form의 분리는 1940년 센서스 이후 long form이 담당하였던 다양한 분석적 중요성을 지닌 인구·가구의 사회경제적 특성과 노동력 상태에 대한 자료의 확보, 나아가 주택센서스로서의 기능을 인구센서스로부터 박탈하였음을 의미하는 것이라고 할 수 있다. 그 결과, 미국의 인구센서스는 다시 원점으로 복귀하여 연방하원 의석배정이나 연방보조금 산정근거가 되는 이른바, 법정인구(法定人口, legal population)를 가능한 한 최고 품질의 수치로 제공하는데 사회적 의미를 부여하는 것으로 방향을 전환하고 있다고 할 수 있을 것이다.

한편 2010년 센서스 본조사의 구성요소로서 자격을 박탈당한 long form에 대해서는, 인구조사국은 American Community Survey라 불리는 독립적인 조사체계로 운영되는 대규모의 월차 단위 표본조사로 대체할 계획을 수립하여 왔다. American Community Survey의 주요 특징들에 대해서는 다음 절에서 논의하게 될 것이다.

2. 가구조사명부 정비작업

세계 각국의 공식적인 국가통계에서 무작위 표본추출을 토대로 한 일부조사로서 표본조사가 도입된 것은 1950년대였으며, 그 후부터 국가통계는 센서스와 표본조사를 양 날개로 하여 체계화를 시도하여 왔다. 표본조사는 잘 알려져 있듯이, 1950년대만 하더라도 단순한 사례조사 이상의 의미를 지니지 않았던 일부조사였지만, 표집이론(標集理論, sampling theory)에 입각하여 표본과 모집단의 관계를 설정함으로써 새로운 형태의 통계조사로 각광받기 시작하였다. 표본조사는 센서스라는 대규모의 전수조사가 지니는 숙명적인 약점인 과다한 조사경비와 속보성의 결여라는 문제를 독창적인 방법으로 극복하고, 모집단의 특성을 반영하는 속보성 자료를 제공함으로써, 센서스만으로는 대응하기 힘든 새로운 속보성 통계자료에 대한 사회적 수요를 충족시키는데 성공하였다. 만약 우리가 아직도 표본조사라는 지적재산을 발굴하는데 성공하지 못하였다면, 오

늘날에도 인구현상은 물론 경제사회 현상의 전반에 대한 시의적절한 통계정보를 센서스보다 상대적으로 저렴한 비용으로 획득하는 것이 사실상 불가능하였다고 할 수 있을 것이다.

표본조사는 그 존립기반이 모집단 정보인데, 이 정보를 제공하는 일차적 원천통계가 센서스였던 것은 두말 할 필요가 없다. 이러한 사실은 표본조사의 도입과 본격적 정착의 과정에서 당연한 일이지는 하겠지만, 센서스의 존립기반에도 반작용을 미치기 시작하였다. 구체적으로, 표본조사가 정당성을 확보하기 위해서는 가능한 한 표본오차가 작은 표본추출을 구체화하는 표본틀(sampling frame)을 정비하는 것이 급선무가 되었다. 센서스 조사구의 명부, 지도의 정비는 센서스의 현장실사는 물론 각종 표본조사를 위한 가구명부의 정비까지 포함하여, 국가통계를 체계적으로 정비하고자 할 때 필요한 통계조사의 기반정보로서도 의미를 동시에 지니고 있었다.

이처럼 조사명부의 정비라는 점에서도, 미국의 2010년 센서스 계획에서 또 하나의 핵심요소로 취급되고 있는 가구명부와 조사구 지도정보의 통합정비는 연방정부통계와 지방통계의 제도개선에 직결된 사업으로서 의미를 지닌다. 현재, 인구조사국이 유지·관리하고 있는 주소파일인 MAF(Master Address File)와 디지털 지도인 TIGER(Topologically Integrated Geographic Encoding and Referencing)에 대해서는 V장과 VI장에서 각각 논의하기로 한다.

V. American Community Survey

본 절에서는 2010년 미국의 인구센서스에서 long form을 대체하게 되어 있으며, 독자적인 조사체계에 의하여 실시되고 있는 American Community Survey 프로젝트에 대하여, 조사계획을 개관하기로 한다.

1. American Community Survey 도입의 역사적 배경

미국의 인구센서스는 연방헌법의 규정에 따라 10년을 주기로, 곧 “Decennial Census”라는 이름으로 실시된다. 이 때문에, 센서스가 실시되지 않는 연도에 대하여, 주 또는 시군구의 지방이나 소수민족 등 하위집단 수준에서 인구의 사회경제적 속성이나 거주관련 특성에 대하여 상세하고도 정확성이 보장된 최고고

품질의 통계자료를 확보하는 것은 제2차 세계대전 이후, 미국의 국가통계체제의 구축과 관련하여 수십 년간에 걸친 정책과제가 되었다. 이에 대한 해결방안으로서 Philip Hauser (1942)의 "annual sample census"라는 연차적 표본센서스, Leslie Kish (1981)의 "rolling sample" 곧 순환표본, Roger Harriot, Bateman and McCarthy (1989)는 "decade census"라는 10년 단위의 순환센서스 등 다양한 제안들이 있었다.

그 중에서 Philip Hauser의 연차적 표본센서스는 제2차 세계대전 기간 중 인적·물적 자원의 동원과 전후의 재조정의 필요성에 따라, 당시 육군성(Army Department), 해군성(Navy Department), 전시생산국(Wartime Production Board) 등 연방정부 전쟁수행기관의 최신 통계자료에 대한 빈번한 수요에 대응하기 위하여, 긴급대처방식의 조사로서 구상된 것이었다. 이것은 인구의 사회경제적 특성이나 주택 등에 관한 조사사항을 부분적으로 순환방식으로 체계화하게 되면, 매년 실시하는 표본조사를 통하여 그 자료를 수집할 수 있을 것이라는 것으로서, 그 자체가 대단히 참신한 발상을 지닌 조사계획이었다. 그러나 표본조사의 모집단인 인구총수에 대해서는 종전의 센서스와 그 후의 인구동태통계, 국내외 인구가동통계를 이용하는 인구분석의 추정방법을 이용할 수 없다는 본질적 문제점을 해소하지는 못하였다. 또 매년 서로 다른 주에 대하여 순환방식의 조사를 실시한다고 하는 Harriot, Bateman, and McCarthy (1989)의 이른바 '10년 단위의 순환센서스'는 주별로 인구집계의 연도가 크게 달라지기 때문에, 이것이 동시성(同時性, simultaneity)을 기본원칙으로 하는 센서스를 대신하는데 어려움이 있다고 판단하여, 인구조사국은 그의 제안을 적극적으로 고려하지 않았다.

American Community Survey (ACS)의 방법론적 기초를 이루는 순환표본의 아이디어는 Leslie Kish가 1970년대 말부터 일련의 논문⁶⁾에서 발표한 "연속측정"(continuous measurement)의 개념에서 출발한다(Alexander, 2000, 2001). Leslie Kish는 1941년에 이미 '연차적 표본센서스'의 아이디어를 친구인 Phillip Hauser에게 소개하였으며, 1981년에는 미국의 경제활동인구조사에 해당하는 Current Population Survey (CPS)를 순환표본조사로 실시할 수 있는 개연성을 포함하여, 소지역에 대하여 상대적으로 빈도가 높은 통계자료 수요에 대하여

6) 순환센서스의 원리를 이론적으로 설명하는 논문은 희귀한 편이다. Leslie Kish(1987, 1990, 1999)가 표본조사보다는 센서스를 특별히 언급하면서 대안적인 표본설계를 언급한 적이 있다. 그는 1999년 논문에서 "인구수를 단순히 헤아리는 것이 아니라, 다양한 사회경제적 변수에 대하여 연차적으로 상세한 통계를 작성하는 일이야말로 많은 나라에서 순환센서스의 주목적이 될 것이다."라고 하고 있다. 그는 많은 논문에서 순환 표본설계에 대하여 언급하였으며, 그것이 바로 미국의 American Community Survey나 프랑스의 rolling census의 기초이론을 제공하는데 기여하였다고 할 수 있다.

적극적으로 대응할 것을 주문하였다. 센서스 본조사의 long form을 순환표본조사로 대체하기로 하는 미국 인구조사국의 American Community Survey 프로젝트에 직접적 계기가 된 것은 Leslie Kish (1990)의 논문 “순환표본과 센서스 조사방법론” (Rolling Sample and Census Survey Methodology)이었다. 1992년 이후, 인구조사국은 순환표본조사를 센서스 본조사의 long form의 대체후보로서 간주하고, 2000년 센서스를 사정권으로 하여 연구성과를 꾸준히 축적하여 왔다. 이 과정에서, Leslie Kish가 제안한 “연속측정”의 개념은 American Community Survey라 불리는 대규모의 표본조사로 발전하였다. 어떤 의미에서 본다면, American Community Survey 프로젝트는 방법론적 구상에서부터 50년 이상의 세월이 흐르면서 현재 도입의 최종단계에 이르렀다고 말할 수 있을 것이다.

American Community Survey 프로젝트를 향한 준비작업이 본격적으로 개시된 것은 1996년이었으며, 그 해 인구조사국은 American Community Survey의 예비조사를 실시하기로 하였다. 1997년, 플로리다 주 Broward County 등 미국 전역에서 4개 군을 선택하여 예비조사를 실시하고, 그 후 예비조사의 대상지역을 점차적으로 확대하여, 1999~2001년에는 미국 전역에서 36개 군을 선택하여, 조사의 실시상황에 관한 정보를 수집하였다. 또 2000년 센서스 계획의 일환으로서, American Community Survey의 조사표와 조사계획에 따라서, 총 1,203개 군에서 C2SS (Census 2000 Supplementary Survey)라 불리는 보충조사가 실시되었다. 이처럼, 복수년도에 걸친 일련의 조사결과를 바탕으로 American Community Survey의 연차별 추계치의 안정성과 이용가능성에 대한 검증작업을 실시하였다(Census Bureau, 2003, 2004, 2006). 또 센서스 long form과의 대체가능성을 검증하기 위하여, 복수년도에 걸친 조사결과에 대해서는 2000년 센서스 long form의 집계결과와 비교연구를 시도하였다(Bennett and Griffin, 2002). 한편, 당초의 예정으로는 American Community Survey 프로젝트가 2003년부터 전면적으로 실시되도록 되어 있었다. 그러나 연방정부의 예산정책 때문에, 2005년 1월부터 매월 25만 가구, 연간 약 300만 가구, 표집률 약 2.5%, 5년간의 총표집률 12.5%로 실시하게 되었으며, 또 2006년 1월부터는 학생기숙사, 요양시설, 교도소, 부랑자 임시수용시설 등 시설인구(施設人口, institutional population)에 대해서도 전면적인 조사가 실시되었다(Kincannon, 2006: 2).

2. 표본추출과 현지실사

American Community Survey 프로젝트는 미국 50개 주, 여기에 미국 연방의 회에 의하여 자치권이 위임된 워싱턴 D.C.와 속주 푸에르토리코를 대상으로 하고 있다⁷⁾. 조사가구는 인구조사국이 보유하고 있는 주소파일 MAF을 이용하여 약 1/480의 표집률로 매월 추출하며, 시설인구는 2.5%의 개인을 조사대상으로 선정하고 있다(Census Bureau, 2003: 7). 또 푸에르토리코에는 약 36,000 가구를 대상으로 하는 지역조사가 실시되었다(Census Bureau, 2003: 12). 다만, 조사결과와 정확성을 전체수준에서 조정하기 위하여, 가구의 표집률은 지역의 인구규모와 특성, 사회인구집단의 규모에 따라 독자적으로 설정하였다. 1,200개미만의 가구가 있는 시군구에 대해서는, 종전에 실시된 시험조사나 2000년, 2001년, 2002년에 실시된 보충조사에서 특히 저조한 응답률을 보인 지역이나 소수 민족 거주지역에 대해서는 표집률을 다른 지역보다 높게 설정하였다⁸⁾(Census Bureau, 2003: 12, 2006: Chap 10-13). 다음으로, 응답부담을 평준화하기 위하여 American Community Survey 프로젝트에서 한번 조사대상으로 선정된 가구는 이른바 "조사휴일제"(Survey Holiday Program)을 두어서 향후 5년 동안은 표집대상에서 제외되도록 하고 있다.

현지실사는 MAF의 주소정보를 바탕으로 우편으로 실시되고 있다. 가구번호, 가로이름, 우편번호⁹⁾ 또는 완전한 지방도로 이름, BOX 번호, 우편번호를 갖는 가구만이 조사표의 발송대상이 된다. 반면, 사서함(PO Box)나 여타 비도시형 가구표시의 주소정보에 대해서는 조사표 발송대상에서 제외된다(Census Bureau, 2003: 20).

먼저, 조사 초대장이 대상가구에 발송되고, 계속해서 인식용 바코드를 인쇄한 조사표가 발송된다. 응답을 완료한 조사표는 인디애나 주 잭슨빌에 소재하고 있는 인구조사국 연방자료처리센터(Federal Processing Center)로 반송하도록 되어 있으며, 3주 후에도 조사표가 반송되지 않는 경우에는 독촉장과 함께 조사표가 다시 발송된다. 다음으로 조사표가 최초로 발송되고 나서 6주가 지나도

7) ACS 프로젝트에서 조사설계 과정은 미국을 주 수준에서 5개 지역으로 구분하고, 5년마다 순환하여 순환형 표본조사를 실시하는 계획이 검토대상이 되었다. 그러나 주 수준에서 조사시점에서 최대 4년의 시간지체가 발생하기 때문에, 이 조사방식은 최종적으로 채택되지 않았다.

8) 200 미만의 가구가 있는 행정단위에 대해서는 센서스의 long form과 마찬가지로, 매년 전가구의 10%를 표집하여, 5년간 총 50%의 가구에 대하여 ACS를 실시하도록 하고 있다.

9) 미국의 우편번호(Zip Code)는 다섯 자리로 이루어지는데, 앞의 세 자리는 주와 도시, 나머지 두 자리는 우편구에 대응하도록 되어 있다.

조사표를 반송하지 않는 경우에는, 현지판매용 전화번호부를 바탕으로, 컴퓨터전화면접(CATI: Computer Assisted Telephone Interviewing)이 전국 세 지역¹⁰⁾에 있는 인구조사국 콜센터의 직원들에 의하여 주로 야간과 주말에 약 25 일간에 걸쳐서 실시된다(Census Bureau, 2004: 2).¹¹⁾ 다만, 컴퓨터전화면접 기간 중 조사표가 반송되는 경우에는, 다시 전화조사의 대상에서 제외된다. 또 콜센터에는 숙련된 직원이 조사표의 질문에 대한 각종 대응방안을 마련하고 컴퓨터전화면접도 거부하는 가구에 대하여 재조사를 실시한다. 조사개시 후, 10주가 경과하였는데도 조사에 불응하는 경우, 조사불응 가구의 1/3을 무작위로 표집하여 조사원이 해당가구를 직접 방문하여 조사를 실시한다.

이와 같이, American Community Survey 프로젝트는 1회의 월차 표본조사에 대하여, 최장 10주간의 조사기간을 설정하고 있지만, 조사기간을 넘어서 제출된 우편조사표는 접수하지 않는다(Census Bureau, 2003, 2004, 2006). American Community Survey 프로젝트에 대해서는, 미국법률 제13편 제141조와 제193조에 의하여 조사표에 대한 응답이 의무화되어 있다(Census Bureau, 2004: 2).

3. Long form과 American Community Survey

미국의 인구센서스는 “Decennial Census”로서, 연방헌법에 따라 10년마다 한 번씩 실시되는 주기적 조사이기 때문에, long form의 조사항목에 대해서도 10년마다 한 번씩 스냅사진과 같은 결과수치를 얻게 된다(Citro and Kalton, 2007). 한편, American Community Survey에서는 대상지역을 인구규모에 따라 3종류로 구분하고 있으며, 유형 A는 매년 조사결과를, 유형 B와 유형 C에 대해서는 연차별로 평활화한 통합자료(pooled data)를 이동평균(移動平均)으로 매년 갱신하고, 추계에 의하여 American Community Survey에서 long form에 상당하는 자료를 매년 확보가능하도록 하고 있다. 사실, 인구 250,000명 이상의 지역에 대해서는 2003년부터 American Community Survey에 의한 연차별 추계결과가 인구조사국 웹사이트 American Factfinder에 업로드되어 있다. 또 인구 65,000~250,000명의 지역에 대해서는 2006년 이후, 인구 20,000~65,000명의 지역에 대해서도 3개년 통합자료가 이용가능해지는 2008년 이후, 마지막으로

10) 인디애나 주의 잭슨빌(Jacksonville)에 있는 연방자료처리센터 부속 콜센터 외에, 애리조나 주의 투스콘(Tucson), 메릴랜드주의 헤저스타운(Hagerstown)에 각각 콜센터가 운영되고 있다.

11) 전화면접을 통한 조사는 7회에 걸쳐서 시도되고, 그래도 조사대상자와 통화가 되지 않는 경우에는 다른 정보원을 통하여 가능성이 있는 번호를 탐색하여 전화접촉을 시도하여 보게 되어 있다(Census Bureau, 2003: 24).

센서스 조사구(약 4000명)와 센서스 블록(통상 600~3,000명)에 대해서도 5년 차의 통합자료를 이용할 수 있는 2010년에는 이용가능해질 것이다. 가령, 2010년 센서스에서 long form에 의한 조사를 실시한다고 하더라도, 통상적인 공표일정으로 본다면 그 결과가 예정되는 것은 2012년이 될 것이기 때문에, 이보다 2년이나 빠른 2010년에 미국 전역에 대한 통계자료를 얻을 수 있으며, 2010년 이후에는 매년 이러한 통계자료가 이용가능하다는 것을 의미하게 되는 것이다 (Census Bureau, 2003, 2004, 2006; Kineannon, 2006)

미국의 센서스는 상주주의(de jure principle)를 원칙으로 인구를 집계하여 왔다. 이 때문에, long form도 상주주의에 따라서 실시되어 왔다. American Community Survey는 지금까지의 long form과는 달리, 기본적으로 상주주의 원칙을 완화하여 현주주의(de facto principle)와의 타협점을 모색하고 하고 있다고 볼 수 있다. American Community Survey 프로젝트의 인구집계에는 2개월 원칙이 적용되어, 2개월 이상 거주하는 가구를 상주지로 간주한다. 계절노동자 중에는 2개월 이상 같은 거주지에서 일을 하고, 계절적으로 국내이동을 반복하는 경우도 적지 않다. 이런 종류의 계절이동에 대해서 American Community Survey에서는 조사대상자를 개별 거주지에서 집계하도록 하고 있다. 또 장기휴가로 고향에 있는 학생이나 평일에는 근무지가 가까운 가구에서 통근하고, 주말에는 집에 돌아오는 통근자의 경우에는 예외적으로 기숙사와 자택을 상주지로 하도록 하고 있다(Census Bureau, 2003: 17-18, 2004, 2006). 이러한 인구의 집계방식의 변경은 American Community Survey가 센서스와는 별도의 표집방법에 의한 월단위 조사를 기준으로 한 연간조사로서 설계되었기 때문에 채택된 것이라고 볼 수 있다. American Community Survey는 집계원칙을 상주주의와 현주주의의 새로운 타협점에서 설정하여 인구를 집계함으로써, 취업 등을 이유로 한 계절이동의 결과 발생하는 지역별 인구이동의 변화를 결과수치에 반영할 수 있도록 하는 구조를 갖추고 있다고 할 수 있다.

4. 회수된 조사표의 처리

회수된 조사표는 접수 후 3주 이내에 입력작업으로 이행하지만, 그 중에서 약 1/3에 대해서는 응답내용에 뭔가 결측치가 포함되어 있다. 이 때문에 콜센터에서는 직원이 Telephone Edit and Follow up (TEF)이라 불리는 전화에 의한 편집 및 추적 작업을 진행한다. TEF는 2000년 센서스에서까지 시도된 적이 없었던 것으로 American Community Survey 프로젝트에서 표본조사의 정

확성을 개선하기 위하여 처음으로 도입된 것이다. 또 American Community Survey의 조사표에는 5명의 가구원만을 기록할 수 있도록 되어 있기 때문에, 6명 이상의 대규모 가구에 대해서는 TEF에 의한 전화면접조사가 이루어지도록 되어 있다(Census Bureau, 2004: 24, 2006).

입력이 종료된 조사자료에 대해서는, 지역의 인구규모에 따라 취급방법이 달라진다. 먼저, 인구 65,000명 이상인 군(郡)에 해당하는 유형 A에 대해서는 2006년 이후, 월차조사 자료를 12개월 통합함으로써 연차자료를 작성한다. 또 인구 25,000명에서 65,000명 미만의 군에 해당하는 유형 B에 대해서는 3년간, 다음은 25,000명 미만의 군에 해당하는 유형 C에 대해서는 5년간의 조사자료를 통합하여 사용한다. 이처럼, 일부의 지역에 대하여 복수연차에 걸쳐 조사자료를 통합하는 것은 인구규모가 작은 유형 B나 유형 C에 해당하는 지역에 대하여, 유형 A와 같은 정도로 결과의 정확성을 획득하기 위하여 필요한 표본수를 확보하기 위하여 이루어지는 절차이다(Census Bureau, 2003, 2004, 2006). 또 유형 B에 대해서는 2008년 이후, 또 유형 C에 대해서도 2010년부터는 통합자료의 이용이 가능해지도록 되어 있다. 이러한 방식으로 2010년부터는 미국의 전체 지역을 대상으로 American Community Survey가 long form을 대신하여 자료를 제공하도록 계획되어 있다.

이외에도 American Community Survey 자료에 대해서는 3단계의 가중보정 작업이 실시되도록 되어 있다(Committee on National Statistics, National Research Council, 2001; Census Bureau, 2006: Chapter 11-13). 제1단계의 “가중복원”(assignment of base weights)은 표집율의 차이와 관련이 있기 때문에, 표집율의 역수로 조사결과를 조정한다. 제2단계의 조정작업은 “회수율에 관련되는 조정”(assignment of CAPI sub-sampling factor)으로서, 회수율의 역수를 제1단계의 조정결과에다 곱하는 조정을 실시하는 것이다. 최후의 조정작업은 “인구의 프로토콜에 의한 조정”(adjustment by population protocols)이라고 불리는 것으로, 센서스 본조사 또는 중간추계에 의하여 얻은 성, 연령, 인종, 종족에 대한 각 분포비율과 제2단계의 조정결과 사이에 조정작업이 실시된다.

5. 결과의 공표

American Community Survey에 대해서는 기본적으로 센서스의 long form과 마찬가지로, 당초에는 전국, 주, 군, 군내의 지역이나 시읍면, 통합도시, 지정센서스지구, 도시통계지역, 선거구 수준에서 집계가 이루어졌지만, 장래에는 센서

스 조사구, 투표구, 미국 인디언 보호구역, 학구, 주의원 선거구, 공개 마이크로 데이터 지역구분, 우편번호 지역, 도시지역, 비(非) 도시지역 등 다양한 지역수준에 대하여 집계범위를 확대하는 것으로 계획으로 되어 있었다(Census Bureau, 2003: 30).

2000년 센서스에서는 인구 10만 명을 넘는 지역구분으로서, 48개의 주와 워싱턴 D.C., 그리고 푸에르토리코에서 공개 마이크로데이터 지역구분(PUMAs)이 설정된 마이크로데이터가 제공되고 있다. 미국 인구조사국은 American Community Survey의 연차별 자료에 대해서도 공개 마이크로데이터의 작성을 계획하고 있다(Census Bureau, 2003: 40, 2004, 2006). 미국의 인구조사법¹²⁾은 개인을 식별할 수 있는 조사자료의 공개를 금지하고 있다. 이 때문에, 공개 제공되는 마이크로데이터에 대하여 수치교환¹³⁾(data swapping), 구분의 통합, 톱코딩(top coding)¹⁴⁾ 등 세 가지의 자료 익명화조치가 실시되고 있다. 또 American Community Survey는 센서스 long form을 대체하는 사실상의 후속 조사이기 때문에, 개인자료의 공개에 대해서는 인구센서스에 필적하는 이른바 “72년 유예조항”(72-year moratorium)을 적용하고 있다.

IV. 가구조사명부의 정비

인구조사국이 센서스 본조사의 long form을 대체하는 조사로서 American Community Survey를 위와 같은 형태로 구상하게 된 것은 미국이 종전에도 가구에 대해서도 우편조사를 광범위하게 실시하였다는 사정이 있었다. 또 앞에서 기술한 American Community Survey의 조사계획이 실효성을 갖기 위해서는 가구조사명부의 존재가 큰 역할을 하였다. 여기서는 인구조사국 지리부(Census Bureau Geography Division)가 정비나 갱신을 책임지고 있는 가구조사명부를 구성하는 2개의 데이터베이스 MAF와 TIGER에 대하여 살펴보기로 한다.

12) 미국연방법률 제13편 제9장 제301조 (a) 항

13) 인구조사국은 가령, 출현빈도가 낮은 특이한 가구에 대해서는 다른 집계지역의 유사가구 사이에 교환이 이루어지고 있다(Census Bureau, 2003: 31; Census Bureau 2006: 13-9).

14) 특정 변수의 경우, 도수분포 상 일정 비율을 넘어서는 모든 사례를 하나의 범주로 통합하여 통계 자료를 정리하는 개인정보 노출을 회피하기 위한 익명화 전략을 말한다(Census Bureau, 2003: 31; Census Bureau 2006: 13-9).

1. 우송용 주소 데이터베이스 MAF

MAF는 인구조사국이 구축한 최초의 영구적 주소 데이터베이스로서 미국 1억 1,500만 가구, 6000만의 기업, 사업장, 각종 시설 등을 포함하는 것이다. 이것은 당초 2000년 센서스에서 사용하기 위하여, 1990년 센서스 시점에서 작성한 주소정보를 바탕으로 하여, 현지조사, 우정사업청(US Postal Service)의 배달순서 파일(DSF) 정보, 여기에 인구조사국의 주소·지역 업데이트 프로그램을 기반으로 하여, 지방정부로부터 제공된 주소정보 등으로 보완함으로써 정비되는 것이다. MAF는 현재 미국 50주, 푸에르토리코 그리고 미국령 도서를 대상지역으로 하여 유지·관리되고 있다.

MAF는 가구명부로서, 조사표 송부용(mail-out) 주소, 도시형 주소에 대해서는 주택번호와 가로 이름, 위도와 경도 좌표, 센서스 조사구 번호, 센서스 블록 번호 등의 센서스 조사지구 정보 및 정보이력에 대한 일련의 정보를 가지지만, 동시에 MAF는 각종 경제조사를 위한 사업장의 소재지 정보도 포함하고 있다(Waite, 2003). 주소는 누락 없이 센서스 블록에 할당되어 있으며, 나중에 언급하는 TIGER(Topologically Integrated Geographic Encoding and Referencing database)와 연계되도록 되어 있다.

MAF는 인구조사국이 주관하는 일련의 가구조사를 위한 명부로 이용되고 있을 뿐만 아니라, 조사결과의 편집이나 복원승수 정보로서 제표작업에서도 이용되고 있다(Census Bureau, 2003: 9, 2004, 2006).

도시지역에 대해서는 MAF의 업데이트에 DSF를 이용할 수가 있다. 이 때문에 MAF가 미국 전체를 통일적으로 망라하는 주소 데이터베이스로 효율성을 유지하기 위해서는, 총인구의 약 15%가 거주하고 국토 전체의 과반을 차지하는 DSF가 사용되지 않는 도시형 가구표시를 가지지 않는 지역의 주소정보를 어떻게 업데이트하는가가 결정적 의미를 지닌다. 이들 지역에 대해서는 조사원의 육안관찰에 의한 주소표시나 거리명칭의 변경 등을 확인하는 작업이 불가피하다. American Community Survey의 조사원이나 인구조사국의 현지조사 요원은 이러한 확인업무를 조사업무의 하위영역으로 설정하고 있다. 인구조사국의 요원들은 실사의 결과 새로운 주소표시와 가로 이름에 대한 최신정보를 ALMI(Automated Listing and Mapping Instrument)라는 소프트웨어를 탑재한 노트북 컴퓨터를 거쳐서 인구조사국에 전달되도록 되어 있다(Census, 2003: 6; Census Bureau, 2006, Chap 9). 이러한 일련의 작업은 지역주소변경시스템(CAUS, Community Address Updating System)로 총칭되는 인구조사국의 프

로그랩에 의하여 일상적으로 업데이트가 이루어지게끔 되어 있다.

2. TIGER

인구조사국에서는 1980년 센서스의 준비과정에서 처음으로 대도시지역을 대상으로 위상기하학에 기초한 도형 데이터베이스 GBF-DIME(Geographic Base File/ Dual Independent Map Encoding)을 구축하였다. 도형 상으로 결절점(結節點, node)을 직선으로 연결하여, 다각형으로 센서스 블록을 표시하였던 이 파일은 주소정보나 지역참조 파일로서는 도움이 될 수 있었지만 디지털지도 작성에는 제대로 대응할 수 없었기 때문에 운영상의 어려움에 봉착하게 되었다.

그 후, 도로의 센터라인에 관한 디지털 정보에 의하여 지역을 구분하였던 디지털 지도가 TIGER로서 인구조사국에 의하여 1990년 센서스 현지실사를 목적으로 독자적으로 개발되었다. TIGER은 ① 거리의 명칭과 센터라인 ②호수, 하천, 수로와 그 명칭 ③ 철도노선 ④지리적 구분과 그 명칭 및 코드(행정구획, 센서스 조사구, 센서스 블록 등) ⑤ 가구의 위치(특정지역) ⑥ 지역을 눈으로 확인가능한 대상물(공항, 학교 등) ⑦ 우편번호와 주소표시(도시형 주소지구)에 관한 정보를 가지고, MAF 주소지 정보가 센서스 블록 수준에서 대응될 수 있도록 하였다. 또 GBF-DIME이 미국 전역의 2% 밖에 담당하지 못하던 데하여, TIGER에서는 그 대상지역이 MAF와 마찬가지로 전 50주, 그리고 푸에르토리코와 미국 속령 도서로 확대되었다. 또 TIGER는 인구센서스만이 아니라, 연도별 인구추계나 5년마다 실시되는 경제센서스의 현지실사에도 사용되고 있다.

TIGER는 연방정부, 주정부, 지방정부가 보유하고 있는 지리정보를 정보원으로 하여, 그것을 민간이 보유하는 항공측량도나 GPS에 의하여 보정함으로써 작성, 업데이트된다. TIGER는 정부만이 아니라 넓게는 민간기관도 함께 사용하고 있다. 이 때문에 인구조사국에서는 일반의 행정기관 등 특히 그 사용에 제한이 설정되어 있는 정보에 대해서는, 데이터베이스의 유지 업데이트를 실시하지 않고 있다(Smith, 2008).

현재의 TIGER은 GPS과 연계되어 고도의 지리정보기능을 갖추고 있다. 이것은 조사결과와 지도표시라고 하는 집계처리결과와 공표에서 효율성을 발휘할 뿐만 아니라, 조사의 실시 면에서 조사대상의 발견을 지원함으로써, 조사원의 현지실사에 따른 작업량을 경감할 뿐만 아니라 집계누락이나 집계중복을 방지하여 집계의 정확성을 향상시키는데 크게 공헌하고 있다(Alexander, 2000: 57, Smith, 2008).

3. MAF/TIGER 개선계획

2000년 센서스는 다른 센서스와 비교하여 상당히 성공적으로 조사를 끝마쳤다고 할 수 있다. 여기에서도 센서스 모니터링 부서의 연방의회 최종보고서에 의하면, 310만 명의 중복조사가 있었고 조사누락도 640만에 이르러, 모두 330만 명이 과소집계된 것으로 파악되고 있다(Brown, 2006: 1). 그러나 상업용의 지도정보시스템 GIS의 본격적 서비스 개시 이전에 인구조사국이 개발하였던 TIGER는 조사대상의 집계지원 장치로서나 데이터베이스로서도 많은 개선의 여지가 있다. 특히 후자에 대해서는 TIGER는 독자적인 데이터 포맷을 가지고, 지점정보의 접속에는 특별한 인터페이스를 필요로 하였다. 이 때문에, 일반의 상용 소프트웨어의 개발업자가 인구조사국의 지도정보를 판독하기 위해서 새로운 소프트웨어를 준비하여야 할 필요가 있었다.

잘못 실시된 센서스 결과는 연방하원 의석수 산정이나 각종 보조금 프로그램의 근거를 위태롭게 할 수 있다는 인식 때문에, 센서스에서 조사표를 반송하지 않는 가구의 정확한 지도상의 특징과 American Community Survey를 위한 지도·주소정보의 업데이트를 위하여 MAF/TIGER 개선계획(MAF/TIGER Enhancement Program)을 인구조사국에 명령하고, 이를 위한 비용으로서 2008년까지 총 5억 달러의 연방예산을 투입할 것을 결정하였다.

개선계획에서는 2010년 센서스 실시를 목표로 하여 ① 주소·가로 위치의 정확성 개선, ② 현대적 데이터베이스 정보의 처리환경의 정비, ③ 지리상 협력관계의 확대강화, ④ 지역주소업데이트시스템(CAUS)의 구축, ⑤ 정기적 평가의 실시와 품질측정의 확대 등 5개의 주요 검토과제를 설정하고, 인구조사국에서는 다른 여타 연방기관¹⁵⁾, 지방정부, 민간기업¹⁶⁾과 제휴하여, 이 작업을 진행하고 있다.

인구조사국은 GIS의 기능을 TIGER에 갖추도록 함으로 해서, 이용 면에서

15) 가령, 미국 교통부(Department of Transportation)가 관리하는 연방고속도로 코드체계(Federal Highway Code System)는 센서스 실시와는 아무런 관계가 없다. 이처럼 센서스 현지실사와 직접 관계가 없는 지도정보도 TIGER에 의한 데이터베이스화의 대상이 되며, 변경사항이 발생하는 경우에는 모두 파일에 반영하여야 한다.

16) Harris Corporation은 2002년에 MAF/TIGER 개선계획 중에서 가로의 중심선의 정확한 지형표시(이 회사는 군의 모든 가로를 실제 주행함으로써 정확한 중심선 정보를 획득하고 있다)(Brown, 2006: 4), 개선된 도로 자료와 각종 지점특성의 결합, 그리고 불충분한 지도표시지구의 정보수집에 관련되는 작업을, 총액 2억 달러로 인구조사국에서 수주하였다. 작업은 이미 2003년에 개시되었으며, 2008년 초에는 완료될 것으로 예상되었다. 그러나 2008년 3월 현재 인구조사국과 Harris Corporation 사이의 커뮤니케이션 실패로 이 작업의 성과가 2010년 센서스에서 제대로 이용될 가능성은 거의 없고, 인구 60만 명의 임시노동자를 센서스 조사원으로 활용하는 프로그램을 구상 중이다.

개선을 시도하였다. 또 미국법률 제13편의 비밀보호규정에 따라서, 일반의 통계 이용자가 접근하는 TIGER에서 MAF를 의도적으로 분리하여 지금까지 별도의 데이터베이스로 유지함으로 해서 비밀보호를 이룩하도록 하였다. 그러나 최근의 데이터베이스 기술의 발전을 토대로 하여 TIGER와 통합되어도 비밀유지가 가능하기 때문에, 양자를 통합 데이터베이스화하는 것이 MAF/TIGER에 있어서 중요한 개선과제가 되고 있다.

종래, 군 단위이기는 하지만, 상호 독립의 데이터베이스로 유지, 업데이트되어 왔던 MAF와 TIGER는 새로운 시스템으로의 이행과 함께, Oracle에 의해서 전국 일체의 통합관리 데이터베이스로 재편되어 왔다. 새로운 시스템에서는 표준 GIS 도구가 상용되는 등, 데이터베이스에 대한 접근과 데이터베이스로서의 기능성이 적극적으로 개선되고 있다. 다음에는 2010년 센서스에 사용이 예정되어 있는 모바일 컴퓨터에 의하여 데이터 교환이 한층 효율적으로 되면서, 새로운 업데이트 정보를 실시간으로 데이터베이스에 반영하는 것이 가능해질 것이라고 기대해 볼 수 있었다. 그러나 2008년 현재 인구조사국이 지난 4년간에 걸쳐서 6억 달러를 Harris Corporation에 투자한 현지실사 데이터 자동화 프로그램(FDCA Program)¹⁷⁾에서 모바일 컴퓨터의 개발실패는 60만 명의 비정규직 센서스 조사원의 고용을 불가피하게 하고 있으며, 이로 말미암아 30억 달러의 추가재원이 소요될 것이라는 예상과 더불어 인구집계의 부정확성에 대한 우려감은 증대하고 있다(New York Times, 2008a, 2008b).

Ⅶ. 요약과 결론

2000년 센서스 실시연도의 바로 2년 전에 내려진 연방대법원의 위헌판결로 인구조사국은 센서스 본조사의 결과를 사후조사 등의 정보를 이용하여 보정함으로써, 정확성을 개선하고자 하는 “Accuracy and Coverage Evaluation Program”의 포기를 불가피하게 하였다. 이 때문에, 인구조사국은 센서스 본조

17) 2010년 센서스에서, 인구조사국은 현지에서 획득한 정보를 직접 캡처하는 자동화된 시스템을 사용하여, 자료처리 비용을 축소하고, 운영상 효율성과 집계의 정확성을 개선하고자 하였다. FDCA 프로그램은 인구조사국 컴퓨터시스템에 현지실시간으로 현지자료를 전달하여 다른 자료와의 통합을 가능하게 하는 최첨단 자동화 시스템으로 기획된 것이었다. 이 시스템은 인구조사국과 센서스 지역사무소의 센서스 조사원의 자료수집활동을 지원하고, 2010년 센서스에서 본격적으로 사용될 수 있도록 하는 것을 최종목표로 선정하였다.

사에서의 인구집계에 대하여 지금까지보다 훨씬 더 정확성이 높은 센서스 자료를 확보하는데 관심을 기울이게 되었다.

2010년 센서스 계획의 최대 특징은 1940년 이후 실시되는 long form에 의한 조사를 본조사에서 분리하여 American Community Survey로 실시하고, 센서스를 short form으로만 실시한다는 것이다. 이로써, 미국의 센서스는 1940년 이전의 센서스 형태로 회귀하는 것이 된다. 센서스를 short form만으로 실시한다는 것은 현지실사에서 long form에 투입되던 인적·물적자원을 short form에 집중투입하고, 결과적으로 센서스 본조사의 집계범위를 개선하는데 기여하게 될 것으로 기대하고 있다.

그러나 이 글에서 이미 지적한 것처럼, 사실경과로서, 표본조사를 이용한 조사결과의 보정에 의한 센서스 집계범위 개선방안이 위헌이라고 판단되어, 그 결과 이에 대한 임시대응조치로서 센서스 집계범위 개선방안의 하나로서 long form의 분리 안(案)이 부상하여 2010년 센서스 계획에 포함되었다고 생각해서는 안 된다는 것이다. 그 후 American Community Survey로 구체화된 순환형의 대형 표본조사로 long form을 대체한다고 하는 조사방식의 아이디어 자체는 1940년대 초로 거슬러 올라간다.¹⁸⁾ 인구조사국은 센서스의 long form에서 얻은 정보를 10년마다 한 번씩이 아니라 훨씬 자주 확보하기 위하여, 당초 센서스의 중간년도에 해당하는 1985년에 5년차 센서스(mid-decade census)를 실시하기로 하는 계획을 수립하였다. 그러나 5년차 센서스는 연방행정예산청의 동의를 얻지 못하고, 결국 미국에서는 한국과 같이 5자로 끝나는 연도에 실시되는 5년차 센서스에 의한 통계정보를 획득한다는 당초의 계획은 완전히 좌절을 맛보게 되었다(Alexander, 2000: 56; Alexander, 2001: 3). 패널의 Recommendation 6.2와 관련하여 이미 언급한 것처럼, 과학원학술평의회 산하에 설치된 패널의 심의과정에서 인구조사국의 대규모의 순환 표본조사 계획이 검토대상에 포함되어 있었다. 이것은 인구조사국이 표본조사 계획의 아이디어를 공표한 직후로서, American Community Survey라고 하는 대규모 표본조사에 의하여 long form 정보의 획득으로 정책방침의 대전환을 구체화하기 위한 준비작업을 개시하였다는 사실을 뒷받침해주고 있다.

1996년에 인구조사국은 American Community Survey를 구체화하기 위한 본

18) 미국에서 지역이나 소수민족에 대해서, 센서스만으로 얻을 수 없는 정보를 정규적으로 확보해야 한다는 논의는 적어도 1941년의 "annual sample census"의 제안으로 소급할 수 있다(Alexander, 2000: 53). 이런 의미에서 American Community Survey의 원형은 annual sample census라고 해야 할 것이다.

격적 작업에 착수하였으며, 수차에 걸친 시험조사에서 조사실시나 표본설계 등에 관련된 조사정보를 수집하였다. 그 중에서 특별히 주목할 만한 것은 2000년 센서스의 long form의 결과와 연속추정에 의한 누적자료를 포함하는 American Community Survey의 결과를 비교함으로써, long form을 American Community Survey로 대체하는데 따라 발생하는 문제점을 상세히 검토하기 시작하였다. 이것은 인구조사국이 당초부터 2010년 도입을 목표로, 센서스 long form을 American Community Survey로 대체하기 위하여 착실하게 준비를 하였으며, 이 과정에서 2000년 센서스의 long form도 그 유효성의 검증수단으로 자리매김 되었음을 말하여 주는 것이다. 이러한 의미에서, American Community Survey는 2000년 센서스의 성패여부에 관계없이, 인구조사국이 당초부터 2010년을 목표로 주도면밀하게 도입계획을 준비하여 온 것이라고 할 수 있다.

이처럼, American Community Survey는 무엇보다도 센서스 본조사의 long form이 지금까지 해온 기능을 대체할 목적으로 도입된 것이다. American Community Survey에 대해서는, long form에 대응하는 조사체계상의 지위를 부여하여, short form와 마찬가지로 조사대상자에게 응답의무를 부과하고 있다. 어떤 의미에서 센서스 본조사에서 분리되었다고 하는 American Community Survey는 센서스 long form의 “분신”(incarnation)으로서, 1940년 이후 센서스가 해온 기능을 센서스 실시방식과는 완전히 다른 형태로 재정립한 것이라고 볼 수 있을 것이다.

그렇다고 American Community Survey가 단순히 long form을 대체하는 기능에만 멈추어 버렸다고 생각해서는 안 될 것이다. 그 이유는 센서스 long form이 10년마다 통계자료를 제공하는데 반하여, 일부 지역에 대해서는 3년 또는 5년간의 조사자료를 통합하여 이것을 이동평균법으로 평활화(平滑化 smoothing)함으로써, 센서스를 실시하지 않는 연도에 대해서도 추정수치로서 연차별 통계자료를 제공할 수 있기 때문이다.

인구조사국은 지역조사의 본격도입으로, 한편으로 센서스의 본조사를 short form으로 특화하여 인구집계의 품질을 개선하는데 전력을 경주하면서, 다른 한편으로 센서스에 요구되는 동시성이라는 요건을 만족시키지 못하기 때문에, 우정사업청이 유지·관리하는 MAF라는 조사명부의 지원을 통하여 독자적인 조사계획에 의하여 지금까지 10년마다 얻을 수 있었던 long form 정보를 매년 업데이트하여 획득할 수 있는 조사체제를 구축하였다고 할 수 있다. 이러한 의미에서 American Community Survey의 도입과 센서스 long form의 폐기처분은, 미국의 센서스 발전사에서 센서스의 원점복귀라는 단순한 의미를 훨씬 넘어서

는 부분이 있다는 사실을 잊어버려서는 안 될 것이다.

Short form으로 특화되고 있는 2010년의 센서스, American Community Survey, 그리고 여타 가구조사도 포함하여, 이들 조사의 성패여부를 결정하는 조사실시의 기반정비로서 자리매김되고 있는 것이 MAF/TIGER의 개선계획이다. MAF/TIGER의 개선은 지역조사의 도입과 함께 2010년 센서스 계획의 양대 지주였다. 이것은 당초 2010년 센서스 본조사를 성공적으로 완료하기 위한 필요조건의 하나이다. 미국 과학원 학술평의회 국가통계위원회 패널보고서가 그 의미를 적극적으로 평가한데서도 알 수 있는 것처럼, MAF/TIGER의 정비가 갖는 통계작성 면에서의 파급효과는 2010년 센서스 실시에만 필수적으로 한정되어서는 안 될 것이다. 이것은 인구조사국만이 아니라 여타 연방행정기관, 주정부, 지방정부 또는 민간기관까지도 망라하여 미국의 각종 조사실시의 기반정비작업이라는 실천적 의미를 지닌다. 여기서는 미국 우정사업본부가 보유하는 주소정보, 그리고 연방정부, 주정부, 지방정부가 보유하는 각종의 통계정보가 동원되고, 데이터베이스가 일상적으로 업데이트되도록 하고 있다. 인구조사국이 사용하는 것을 주목적으로 하고 있지만, 종전처럼 센서스 실시시점에서 현지실사의 필요성 때문에 정비되는 센서스 조사구 정보로서가 아니라, 기존의 행정정보(우정사업본부의 등록정보)를 베이스로 하는 파일이 독자적으로 정비되고 있다는 의미에서, 프레임 정비기능은 센서스에서 상대적으로 자립화하는 움직임을 MAF/TIGER 정비사업에서 읽어볼 수 있을 것이다.

미국의 2010년 센서스는 얼마 있지 않아서 실시될 것이다. 그러나 2010년 센서스는 적어도 두 가지 점에서 큰 의미를 지닌다고 할 수 있을 것이다. 먼저, 2010년 센서스는 headcount, 곧 사람의 머릿수를 헤아리는 전통적 센서스로 복귀하게 된다. 그러나 그것으로 2010년 센서스의 의미는 상실되지 않는다. 지금까지 long form으로서 센서스가 해왔던 사회적 기능의 일단을 American Community Survey라는 새로운 패러다임의 조사로서 재구축하고, 가구조사명부의 정비기능을 센서스에서 분리하여 새로운 발전의 양상을 보이고 있는 것에 주목해야 할 필요가 있을 것이다. 그럼에도 불구하고, 현재의 판단으로는 미국의 2010년 인구센서스가 인구조사국이 개발하고자 했던 모바일 컴퓨터의 문제점으로 인하여, 인구집계가 하이테크를 극대화한 방식으로 실시되지 못하고, 60만 명의 비정규직 센서스 조사원을 고용해야 한다는 새로운 도전에 직면하면서, 전통적 센서스의 문제점이 short form과 long form의 완전분리, 그리고 short form만의 전수조사 실시라는 혁신적 센서스 개혁으로도 완전히 극복되지 못할 수 있다는 딜레마가 있음을 직시할 필요도 있을 것이다(New York Times, 2008a, 2008b).

참고문헌

- 전광희 (2008) “영국의 2001년 One Number Census 프로젝트” 통계청통계개발원 보고서 2005년 인구주택총조사 분석보고서.
- Alexander, C. H. (2000) “The American Community Survey and the 2010 U. S. Census” A paper presented at INSEE-Eurostat seminar on census after 2001 (Paris, November 2000).
- _____ (2001) “Still Rolling: Leslie Kish’s “Rolling Samples” and the American Community Survey” Proceedings of Statistics Canada Symposium 2001, Achieving Data Quality in a Statistical Agency: A Methodological Perspective.
- Anderson, Margo A. (1988) *The American Census: A Social History*. New Haven: Yale University Press.
- Bennett, C. H. and D. Griffin (2002) “Race and Hispanic Origin Data: A Comparison of Results from the Census 2000 Supplementary Survey and Census 2000” A paper presented at the Joint Statistical Meetings, August 2002.
- Citro, Constance F. and Graham Kalton(eds) (2007) *Using the American Community Survey: Benefits and Challenges, Panel on the Functionality and Usability of Data from the American Community Survey*, Washington D.C.: National Academy Press.
- Committee on National Statistics, National Research Council (2001) *American Community Survey: Summary of a Workshop*, Washington, D.C.: National Academies Press.
- Cork, Daniel L. Michael L.Cohen, and Benjamin F. King (eds) (2004) *Reengineering the 2010 Census: Risks and Challenges, Panel on Research on Future Census Methods*, National Research Council, Washington, D.C.: National Academies Press.
- Galdi, D. (2005) “Spatial Data Storage and Topology in the Redesigned MAF/ TIGER System” Washington, D.C.: Census Bureau. (http://www.census.gov/mtep_obj2/topo_and_data_stor.pdf)
- Hauser, P. M (1942) “Proposed Annual Census of the Population” *Journal of the American Statistical Association* 7.
- Herriot, R.A., D.B. Bateman, and W. F. McCarthy (1989) “The Decade Census Program-A New Approach for Meeting the Nation’s Needs for

- Subnational Data” *Proceedings of the Social Statistics Section*, American Statistical Association.
- Kincannon, C. .H (2006) “Apportionment in the Balance: A Look into the Progress of the 2010 Decennial Census” *A Report given at the Subcommittee on Federalism and the Census*, U.S. House of Representatives on 1st March 2006
- Kish, L (1981) *Using Cumulated Rolling Samples to Integrate Census and Survey Operations of the Census Bureau*, Washington, D.C. Government Printing Office.
- Kish,L (1990) “Rolling Samples and Censuses” *Survey Methodology* 16: 63-79.
- Edomonton, B. and C. Schultze (eds) (1996) *Modernizing the U.S. Census, Panel on Census Requirements in the Year 2000 and Beyond*, Washington, D.C.: National Academies Press.
- New York Times (2008a) “Editorial: The Census at ‘High Risk’” March 25, 2008.
- _____ (2008b) “Editorial: Census Damage Control” June 23, 2008.
- Smith, Alex (2008) “The TIGER Modernization Project: The US Census Bureau Seeks to Update and Improve TIGER Data for the 2010 Census” (http://www.personal.psu.edu/aws157/Project_3_report.htm).
- Trainer, Timothy (2005) “The MAF/TIGER Enhancement Program: The Mechanics and Maintenance of a Large-scale National Spatial Database” Washington, D.C.: Census Bureau. (<http://www.cartesia.org/geodoc/iccc/pdf>).
- U. S. Census Bureau. (2003) “American Community Survey Operations Plan-Release 1” March 2003. Washington, D.C.: Census Bureau. (<http://www.census.gov/acs/www/Downloads/OpsPlanfinal.pdf>).
- _____ (2004) *American Community Survey: A Handbook for State and Local Officials*. Washington, D.C.: Census Bureau. (www.census.gov/acs/www/Downloads/ACS04HSLO.pdf)
- _____ (2006) *American Community Survey: Design and Methodology*. Technical Paper 67. Washington, D.C.: Census Bureau. (<http://www.census.gov/acs/www/Downloads/tp67.pdf>)
- Vitrano, Frank A. (2004) *The Concept and Method of American 2010*

Population Census. Washington, D.C.: Census Bureau.

Waite, Preston J. (2003) "The Reengineered 2010 Census" Proceedings of Statistics Canada Symposium 2003, Challenges in Survey Taking for the Next Decade.

_____ & Birnbaum I. Nicholas. (2002) *Census 2000 Methods and the Vision for the 2010 Census*. Washington, D.C.: Census Bureau.

[2008. 5. 12 접수 | 2008. 8. 24 채택]

Transformation and Future Prospect in the U.S. Census: Focusing on 2000 Census Experience and 2010 Census Plans

Kwang-Hee Jun

This paper discusses transformation of the US population census since 1990 and its future prospects by mentioning the 2000 census experience and the 2010 census plans. First, it examines the recommendations written for the 2000 census by the Committee on National Statistics, National Research Council, such as introduction of statistical estimation, response rate improvement, long form improvement, adjustment in differential undercount by race/ethnicity, alternative census methods, collection of small-area statistics in non-census years, and census frame enhancement, and describes how the US Supreme Court decision of 1998 led the Census Bureau to fail in conducting the US version of One Number Census which uses statistical estimation by matching actual enumerations from the Census main survey and post-enumeration survey. Second, it examines one key element of the 2010 U. S. census, say, the separation of long form from short form and describes the main features of American Community Survey, a rolling census which replaces long-form component of the traditional US "Decennial Census" Another element is MAF/TIGER Enhancement Program which aims to improve enumeration accuracy in the traditional short-form census and help the Census Bureau introduce a mobile computer system as part of high-tech census operation. In this paper, it is pointed out that the separation of long form from short form is not an accidental one which results from the US Supreme Court decision, but the Census Bureau at this time in 2008 worries about the accuracy of enumeration because it has failed to develop a mobile computer system and will have to canvass 115 million households by paper and pencil by hiring 600 thousand temporary census workers.

Key Words: U.S. census, Census Bureau, short form, long form, AF/TIGER, American Community Survey, One Number Census, statistical estimation, mobile computer, rolling census, sample census, alternative census.