연구논문

국내 유선 전화조사에서 표본추출틀의 포함률

Coverage Rates for Households by Landline Telephone Frames in Korea

홍성준* · 박소형** · 김선웅***
Sung - Joon Hong · So - Hyung Park · Sun - Woong Kim

국내 전화조사에서 면접대상인 가구(또는 개인) 표본을 뽑기 위한 표본추출틀로서 전화번호부 또는 이와 대등한 전화번호 데이터베이스가 주로 사용된다. 최근에 전화번호부 등재를 원하지 않거나 이동전화의 사용으로 유선전화를 보유하지 않은 가구들이 늘면서 전화번호부에 포함되지 않은 가구들이 증가하고 있다. 그런데 전화번호부의 가구 포함률이 낮을수록 포함편향으로 인하여 조사의 정확성이 떨어질 수 있다. 따라서 국내 전화조사의 정확성 판단을 위해 전화번호부의 가구 포함률 추정은 매우 중요하다. 본 논문에서는 먼저 국내 가구들의 전화보유 분포를 살펴보고 이를 다른 국가들과 비교한다. 그리고 전화번호부의 특성들을 설명하고 간접적, 직접적 방법으로 전화번호부의 가구 포함률을 추정하며 그 실제 포함률이 전국적으로 65.6%에도 못 미쳐 조사결과의 신뢰성 확보가 어려운 수준임을 보인다. 또한 이러한 비포함 문제를 해결하기 위해서는 효율적 임의번호추출법의 개발이 필수적임을 밝히며 이에 따른 몇 가지 연구방향들도 제시한다.

주제어: 표본추출틀, 포함편향, 비등재 가구, 비포함, 포함률, 임의번호추출법

Landline telephone surveys of the population of households or individuals in Korea often use telephone directories as sampling frames. Recently, the frequency of unlisted numbers in the directories has been increased and the number of households without landline phones has become larger with a spread of mobile phones. Landline telephone coverage has currently reached to a level that raises concerns about the accuracy due to a coverage bias on the statistics in question. In this paper, we first present the distribution of telephone ownership in Korea and make a comparison with some selected countries. Second, we describe the characteristics of telephone directories. Next, we directly or indirectly estimate the telephone coverage rates of the frames, and show that it may nationally be lower than 65.6% based on additional information. We conclude with remarks about future studies to reduce coverage

* 동국대학교 대학원 통계학과 박사과정
** 동국대학교 대학원 통계학과 박사과정
*** 교신저자(corresponding author): 동국대학교 통계학과 부교수 김선웅.
E - mail : sunwk@dongguk.edu

bias, including the developments of efficient random digit dialing sampling methods.

Key words: sampling frames, coverage bias, unlisted households, noncoverage, coverage rate, random digit dialing sampling methods

Ⅰ. 서론

정치 및 선거여론조사, 사회조사, 공공정책 관련 조사 등의 많은 표본조사에서 가구(또는 개인)에 대한 자료수집방법으로서 전화면접이 흔히 사용된다. 이 전화면접은 일반전화인 유선전화를 보유한 가구들을 대상으로 주로 진행된다. 전화면접의 대상이 되는 표본가구들을 추출하기 위해서는 표본추출틀(sampling frames)로 사용될 수 있는 '유선전화를 보유하고 있는 가구들의 목록'이 필요하며 이 목록의 대표적인 예로서 전화번호부를 들 수 있다.

그런데 유선전화를 보유하지 않은 가구들(이하 '유선전화 미보유 가구들'로 표기) 또는 유선전화를 보유하고 있으나 전화번호부에는 비등재된 가구들(이하 '유선전화 비등재 가구들'로 표기)은 전화번호부와 같은 표본추출틀에 포함되지 않아 표본조사의 대상에서 빠지게 된다. 이러한 것을 '가구 비포함 문제(household noncoverage problems)'라 한다.

유선전화 미보유 가구들의 비율(이하 '유선전화 미보유율'로 표기)과 유선전화 비등재 가구들의 비율(이하 '유선전화 비등재율'로 표기)을 통칭으로 비포함률(noncoverage rate)이라 하며 이 비포함률이 높을 때 가구 비포함 문제는 커질 수 있다. 좀더 구체적으로 설명하면 비포함률이 높을 때 즉, 포함률(coverage rate)이 낮을 때 모집단을 구성하는 전체 가구들에 대한 특성치와 표본추출틀에 포함되는 가구들의 특성치 간의 차이를 나타내는 포함편향(coverage bias)이 커질 수 있다. 예를 들어 Blankenship(1977)과 Moberg(1982) 등은 유선전화 등재 가구들과 비등재 가구들간에 성별, 연령, 직업 등을 포함한 다양한 사회인구학적 특성들이 차이가 있다는 것을 밝혔는데 이는 비등재 가구들이 제외된 포함률이 낮은 표본추출틀을 사용할 경우 '조사의 정확성'이 떨어질 수 있다는 것을 의미한다. 따라서 비포함률은 조사품질(survey quality)의 기본적이면서 중요한 척도이다.

유선전화 미보유 가구들 때문에 발생하는 비포함 문제는 과거에 Trewin & Lee(1988)와 Thornberry & Massey (1988)에 의해 심도 있게 다루어졌으며, 특히 Trewin & Lee는 전세계 주요 국가들의 비포함 문제를 비교하였다. 또한 유선전화 비등재 가구로 인한 비포함 문제도 미국 등에서 1960년대부터 많은 학자들에 의해 연구가 되었는데 Lepkowski (1988)는 이 비포함 문제를 해결하기 위해 개발된 임의번호추출법[random digit dialing(RDD) sampling method]를 포함한 다양한 표본추출방법들을 체계적으로 소개하였다.

Kim & Lepkowski(2002a), Kim & Lepkowski(2002b), Kuusela et al.(2008), 그리고 Blumberg et al.(2008)에 의해 논의된 바와 같이 1990년대 말부터는 핀란드, 포르투칼, 오스트리아, 벨기에, 프랑스, 이탈리아, 영국 등 다수의 서유럽 국가들과 미국을 포함한 북미 국가들에서 이동전화의 이용이 급증하였고, 유선전화를 보유하지 않고 대신에 이동전화를 보유하는 가구들(이하 '이동전화만 보유한 가구들'로 표기)이 점차 늘면서 국가에 따라 다소 차이가 있지만 2000년대에는 유선전화 미보유율이 1990년대 보다 전반적으로 증가하였다. 이로 인하여 표본추출틀의 비포함 문제에 대한 관심이 더욱 고조되었고 이 문제를 해결하기 위한 방안들이 여러 국가들에서 모색되고 있다. 이에 대한 내용은 Lepkowski et al.(2008)에 의해 상세히 언급되었다.

한편, 국내 유선전화 비포함 문제는 김선웅(2004)이 국제전기통신연합(International Telecommunications Union; ITU)과 통계청의 보고서 자료를 근거로 다루었다. 그는 국내 이동전화의 급속한 보급에 따라 유선전화 보유가구들의 비율이 감소 추세임을 보여 주었고, 예를 들어 서울특별시의 경우는 이로 인하여 허명회 외(2003, 2004)가 언급한 비포함률인 약 30%보다 더 높은 수준이 될 수 있다는 것을 논의하였다.

이러한 논의와 관련하여 '조사의 정확성 측면에서 표본추출틀이 최소한 어느 정도의 비포함률을 유지해야 하는가?'에 국내 연구자들의 관심이 모아질 수 있다. 사실 Collins & Sykes(1987)가 언급한 것처럼 경험적으로 시장조사(market surveys)에서는 20% 이상의 비포함률이 매우 일반적이지만 사회조사나 공공정책 관련 조사 등에서 이러한 비포함률을 갖는 경우 조사결과를 신뢰하기 어렵다. 더욱이 정치 또는 선거여론조사 등과 같이 정당 또는 유권자들이 첨예의 관심을 갖기 때문에 보다 정확한 결과를 필요로 하는 조사의 경우 훨씬 낮은 비포함률이 요구될 것이다.

따라서 '국내 유선전화 표본추출틀의 실제적인 비포함률은 얼마인가?,' 다시 말해서, '국내에서 유선전화조사를 실시할 때 전국적으로 또는 지역별로 전체 가구들 중 몇 퍼센트

가 전화번호부와 같은 표본추출틀에서 제외되고 과연 나머지 몇 퍼센트가 포함되는가?' 에 대한 연구는 중요하다.

본 논문에서는 먼저 국내 표본조사자료들을 근거로 가구들의 전화보유 분포를 살펴보고 이를 몇 개 주요 국가들과 비교한다. 둘째, 국내의 전화조사에서 표본추출틀로 사용될 수 있는 것들을 정리하고 이들 중 쉽고 편리하게 표본추출틀로 사용할 수 있는 전화번호부인 'SUPERPAGE'[1]의 특성을 확인한다. 또한 국내 표본조사 결과 및 인구주택총조사 결과 등을 이용하여 2005년 말 기준으로 이 전화번호부의 전국 및 지역별 비포함률을 추정하고 그 분포를 살펴본다. 셋째, 이 전화번호부를 이용하여 전국 규모의 조사를 실시할 때 실제 포함률은 65.6%보다 다소 낮은 심각한 수준임을 보인다. 넷째, 비포함 문제의 해결을 위한 외국에서의 과거 연구들을 정리하고 마지막으로 국내 전화조사에서의 비포함 문제를 해결하기 위해 필수적인 임의번호추출법의 개발과 이에 따른 앞으로의 연구방향을 논의한다.

Ⅱ. 본 론

1. 가구들의 전화보유 분포와 최대 포함률

김선웅(2004)에 의해 논의된 것처럼 통계청의 사회통계조사보고서(2000)와, 정보화실태조사보고서(2001, 2002)에 따르면 국내에서 유선전화를 보유한 가구들의 비율(이하 '유선전화 가구보유율'로 표기)이 1997년, 2000년, 2001년, 2002년에 각각 95.2%, 93.4%, 91.8%, 91.6%로서 조금씩 감소한 반면 이동전화를 보유한 가구들의 비율(이하 '이동전화 가구보유율'로 표기)이 2000년, 2001년, 2002년에 각각 77.2%, 79.4%, 83.1%로 점차 증가하였다. 이러한 유선전화 가구보유율의 감소는 이동전화 보급의 증가에 따른 영향이므로 2002년 이후의 유선전화 가구보유율 또는 이동전화 가구보유율의 변화에 관심을 가질 필요가 있다. 그러나 2003년부터는 통계청에서 진행하던 정보화실태조사가 정보통신부(한국인터넷진흥원)로 이관되면서 가구별로 유선전화와 이동전화를 포함한 정보통신기기 보유 여부를 묻는 조사항목이 제외되었고, 통계청의 사회통계조사에서도 정보통신 관련 조사가 중단이 되어 2003년 이후의 전화 가구보유율에 대한 자료를 이용할 수 없다. 따라서 2002년 정보화실태조사가 유선전화와 이동전화 등의 가구보유율을 알 수 있는 가장 최근 조사라 할 수 있다.

---

1) 전화번호부 표지 상의 제목이 한글 '슈퍼페이지'가 아닌 영문 'SUPERPAGE'로 되어 있음.

〈표 1〉 한국, 2002년 전화보유 분포(백분율)

(단위: %)

전화보유 가구유형	유형별 보유율	유선전화 보유율 ①+②	이동전화 보유율 ①+③	전화 보유율 ①+②+③
유선전화와 이동전화 모두 보유	75.3①	91.6	83.1	99.4
유선전화만 보유	16.3②			
이동전화만 보유	7.8③			
유선전화와 이동전화 모두 미보유	0.6④			

이 정보화실태조사는 2002년 3월에 층화확률비례계통추출법(stratified systematic sampling with probability proportional to size measures)으로 추출된 전국의 28,179개 가구를 대상으로 방문면접방식으로 진행되었다. 정보화실태조사보고서 상에서는 유선전화 가구보유율과 이동전화 가구보유율에 대한 정확한 수치가 제공되지만 좀 더 상세한 정보, 예를 들어 이동전화만 보유한 가구들의 비율과 같은 정보는 제공되지 않는다. 그러므로 이 조사의 원자료를 분석하여 〈표 1〉을 얻었다.

앞에서 정보화실태조사보고서(2002)에 따른 유선전화 가구보유율과 이동전화 가구보유율을 각각 91.6%와 83.1%이라고 하였는데, 이들 수치는 원자료를 분석해서 나온 〈표 1〉의 유선전화 가구보유율(①+②)과 이동전화 가구보유율(①+③)과 각각 동일하다. 그리고 유선전화와 이동전화를 모두 보유한 가구들과 유선전화만 보유한 가구들 각각의 비율이 75.3%, 16.3%이며, 유선전화와 이동전화를 모두 가지고 있지 않은 가구들의 비율은 0.6%에 불과하였다. 특히 주목해야 할 점은 이동전화만 보유한 가구들의 비율이 7.8%라는 사실이다.

이러한 현상이 다른 국가에도 나타났는지 살펴볼 필요가 있다. 영국의 전화보유율을 나타내는 〈표 2〉는 영국전화통신국(Office of Telecommunications United Kingdom 2003)이 유선전화와 이동전화 시장에서 소비자 행동을 파악하고자 실시한 조사의 보고서 상에 있는 2002년 2월과 2003년 2월 자료이다.

〈표 1〉과 〈표 2〉를 함께 살펴볼 때 영국의 전화보유 분포는 한국의 전화보유 분포와 비슷하며 특히 한국의 2002년과 영국의 2003년은 거의 동일함을 알 수 있다. 이러한 영국에 대해 2003년 이후의 자료를 살펴보는 것이 유용할 수 있다. 〈표 3〉은 Kuusela et al.(2008)가 제시한 2006년 영국의 전화보유 분포이다.

〈표 2〉 영국, 2002년 및 2003년 전화보유 분포(백분율)

(단위: %)

전화보유 가 유형	유형별 보유율	유선전화 보유율 ①+②	이동전화 보유율 ①+③	전화 보유율 ①+②+③
유선전화와 이동전화 모두 보유	76 / 76①	95 / 92	80 / 83	99 / 99
유선전화만 보유	19 / 16②			
이동전화만 보유	4 / 7③			
유선전화와 이동전화 모두 미보유	1 / 1④			

주: '/'는 '2002년/2003년'을 나타냄

〈표 3〉 영국, 2006년 전화보유 분포(백분율)

(단위: %)

전화보유 가구유형	유형별 보유율	유선전화 보유율 ①+②	이동전화 보유율 ①+③	전화 보유율 ①+②+③
유선전화와 이동전화 모두 보유	80①	90	90	100
유선전화만 보유	10②			
이동전화만 보유	10③			
유선전화와 이동전화 모두 미보유	0④			

〈표 2〉와 〈표 3〉을 함께 살펴볼 때 영국은 2002년, 2003년, 2006년 이동전화만 보유한 가구들의 비율이 각각 4%, 7%, 10%로 그 비율이 점차 증가하였고, 이로 인하여 유선전화 가구보유율은 2002년 95%로부터 2006년 90%로 감소하였다. 또한 2006년 유선전화와 이동전화 모두 미보유인 가구들은 0%이고 이동전화만 보유한 가구들의 비율이 10%이므로 전체 비포함(미보유)률이 10%임을 알 수 있다. 따라서 영국은 2006년 시점에서 유선전화 표본추출틀이 거의 완벽할 경우 전체 가구들 중 최대 90%가 포함될 수 있다는 것을 알 수 있다.

한국의 경우는 앞에서 언급한 것처럼 2003년 이후의 전화보유 분포에 대한 자료가 없어 영국과 같은 유사한 변화가 있었는지 정확히 파악할 수는 없으나, 통계청에서 이관된 정보통신부(2005, 2006)의 정보화실태조사(약 7,000가구를 대상으로 반기별로 실시)에는 정보통신 관련 지출비용을 묻는 조사항목이 있어 이를 이용하여 '이동전화만 보유한 가구들'의 비율을 대략적으로 추정할 수 있다.

〈표 4〉 한국, 이동전화만 보유한 가구들의 비율

(단위: %)

2005년 12월	2006년 6월
8.9	10.6

2005년 12월과 2006년 6월의 두 기준 시점에서의 '이동전화만 보유한 가구들'의 비율을 추정하기 위하여 '유선전화 지출비용'은 '0원'이면서 '이동전화 지출비용'은 '0원'이 아닌 가구들을 '이동전화만 보유한 가구들'로 간주하여 정보화실태조사의 원자료를 분석한 결과 〈표 4〉와 같은 비율들을 얻었다.

〈표 4〉에 나타나는 두 비율들은 〈표 3〉에서 제시된 2006년 영국의 비율 10%와 거의 비슷하다. 따라서 이를 근거로 다음과 같은 추론이 가능하다.

2005년 12월 기준으로 한국에서 유선전화와 이동전화 모두 미보유인 가구들의 비율이 〈표 1〉의 2002년 수준과 동일한 0.6%라 간주하면 〈표 4〉의 2005년 12월의 이동전화만 보유한 가구들의 비율 약 9%와 합쳐 비포함(미보유)률은 약 10%가 된다. 그러므로 2005년 12월 기준으로 한국도 거의 완벽한 유선전화 표본추출틀을 사용할 경우 전국의 가구들 중 최대 약 90%가 포함될 수 있다.

〈표 5〉 핀란드, 2003년 전화 보유 분포(백분율)

(단위: %)

전화보유 가구유형	유형별 보유율	유선전화 보유율 ①+②	이동전화 보유율 ①+③	전화 보유율 ①+②+③
유선전화와 이동전화 모두 보유	58①	65	93	100
유선전화만 보유	7②			
이동전화만 보유	35③			
유선전화와 이동전화 모두 미보유	0④			

〈표 6〉 미국, 2004년 전화 보유 분포(백분율)

(단위: %)

전화보유 가구유형	유형별 보유율	유선전화 보유율 ①+②	이동전화 보유율 ①+③	전화 보유율 ①+②+③
유선전화와 이동전화 모두 보유	46.4①	88.6	52.4	94.6
유선전화만 보유	42.2②			
이동전화만 보유	6.0③			
유선전화와 이동전화 모두 미보유	5.4④			

한국이나 영국과는 다른 전화보유 분포를 갖는 국가들도 있다. 예를 들어, <표 5>와 <표 6>은 Kuusela(2003)와 Tucker(2005)에 의해 각각 제시된 자료를 가지고 작성한 핀란드와 미국의 가구유형별 전화보유율을 보여 준다.

김선웅(2004)에 의해 언급되었던 것처럼 핀란드는 전 세계 국가들 중 이동전화의 영향이 가장 큰 국가들 중의 하나이다. 이는 <표 5>에 나타난 것처럼 유선전화 가구보유율(65%)이 오히려 이동전화 가구보유율(93%)보다 훨씬 낮으며 '이동전화만 보유한 가구들'의 비율이 우리나라의 2002년 비율(7.8%)의 4배가 넘는 35%라는 사실에서 알 수 있다. 그러므로 핀란드는 2003년 비포함(미보유)률이 35%이며, 따라서 거의 모든 유선전화가구들을 포함하는 전화번호부를 표본추출틀로 사용하더라도 전체 가구들 중 약 2/3만이 표본추출의 대상이 되므로 유선전화만 이용한 조사 시 포함편향이 매우 클 수 있다는 것을 알 수 있다.

그리고 <표 6>에서 볼 수 있는 것처럼 2004년 미국의 유선전화와 이동전화를 모두 보유한 가구들의 비율과 유선전화만 보유한 가구들의 비율은 다른 연도이기는 하지만 <표 1>의 2002년 한국과 큰 차이가 있다.

요약컨대, 2005년도 기준으로 우리나라는 이동전화만 보유하거나 유선전화와 이동전화 모두 보유하지 않은 유선전화 미보유 가구들로 인한 비포함률 약 10%에 기인하여 국내의 모든 유선전화가구들을 포함하는 거의 완벽한 표본추출틀을 사용하더라도 가능한 최대 포함률 $P_{\max}$의 수준은 다음 (1)과 같으며 이는 영국과 거의 비슷하다.

$$\text{2005년 12월 최대 포함률 } P_{\max} \approx 0.9(90\%) \qquad (1)$$

2. 표본추출틀의 포함률

한국에서 전화조사를 위한 표본추출틀을 사용할 때 포함률은 유선전화 비등재 가구들로 인하여 식(1)의 약 90% 수준보다 크게 낮다. 이에 대하여 구체적으로 살펴보면 다음과 같다.

1) 표본추출틀

유선전화 보유가구들에 대한 표본추출틀로 사용될 수 있는 것들은 크게 다음과 같이 6가지로 구분할 수 있다.

- 케이티(KT)의 유선전화번호 데이터베이스
- 하나로텔레콤(hanarotelecom)의 유선전화번호 데이터베이스

· 한국전화번호부(KTD)의 유선전화번호 데이터베이스
· 한국인터넷서비스(KOIS)의 유선전화번호 데이터베이스
· 한국인포데이타(KOID)의 유선전화번호 데이터베이스
· 한국전화번호부(KTD)의 유선전화번호부 – SUPERPAGE

KTD, KOIS, KOID의 유선전화번호 데이터베이스는 KT와 하나로텔레콤 등으로부터 제공된다. KTD는 국내 유일의 전화번호부 업체이고 KOIS와 KOID는 114번호 안내서비스 등을 운영하며 대전 이북과 대전을 포함한 이남 지역을 각각 사업권역으로 하고 있다. KT, 하나로텔레콤, KTD, KOIS, KOID의 유선전화번호 데이터베이스는 실시간으로 갱신된다. 반면에 KTD의 SUPERPAGE는 KTD 유선전화번호 데이터베이스를 편집하여 발간되는 전화번호부이므로 갱신되는 데 상당한 시간이 걸린다.

그런데 이들 유선전화 데이터베이스는 실시간으로 갱신되어 표본추출틀로서의 유용성이 큼에도 불구하고 최근 강력하게 시행되고 있는 '개인정보보호법'으로 인하여 KOIS 등이 일부 조사업계에 제한적으로 텍스트 형태로 판매하고 있으나 실시간으로 제공되지는 않으며 구입비용도 매우 높아 일반 연구자들의 이용이 어렵다.

반면에 KTD의 유선전화번호부인 SUPERPAGE의 '인명편'이 일반에 공개되고 있어 가구들에 대한 표본추출틀로 저렴하고 쉽게 이용할 수 있다. 따라서 SUPERPAGE의 특성들을 살펴볼 필요가 있다.

2) SUPERPAGE의 특성

(1) SUPERPAGE의 구성

SUPERPAGE는 국내의 유일한 유선전화번호부이다. 이 전화번호부는 시 지역은 2년 주기로, 군 지역은 1년 주기로 발행이 되며 일부 지역은 시·군단위들이 전화번호부 한 권으로 묶여서 발행된다.

대도시 지역의 경우는 몇 개의 구들을 전화번호부 한 권으로 묶어서 발행한다. 예를 들어 서울의 경우 총 8권의 전화번호부가 발행되며 각 권별로 몇 개의 구들이 묶여 있다. 예로서 강서구, 구로구, 양천구는 한 권으로 묶여 발간되며 각 구의 시각적 구분이 가능하다.

SUPERPAGE 전화번호부는 총 173권으로 구성되며 이들을 시군지역별로 나누면 총 232권의 분량이 된다.

전화번호부 상에는 전화서비스의 가입자 성명, 주소, 전화번호가 수록되어 있고 성명의 가나다순으로 나열되어 있다. 또한 수록기준일이 기재되어 있어 KTD의 데이터베이스를

이용하여 전화번호부를 작성한 시점을 정확히 알 수 있다.

전화번호부는 현재 국내 전화조사에서 가구들에 대한 표본추출틀로 널리 사용되고 있으며 Lepkowski(1988)가 논한 방법들을 이용하여 표본을 추출할 수 있다.

(2) SUPERPAGE의 포함률 추정

SUPERPAGE에는 공개를 원하지 않는 가입자들의 정보는 수록되지 않으므로 비등재 가구들이 발생한다. 따라서 SUPERPAGE의 비등재율을 파악할 수 있으면 포함률을 알 수 있다.

SUPERPAGE의 수록기준일이 시·군 지역별로 다르므로 다음과 같은 2가지 방법, 즉 간접적인 방법과 직접적인 방법에 의해 SUPERPAGE의 포함률을 추정할 수 있다.

① 간접적 방법에 의한 포함률 추정

SUPERPAGE는 지역별로 수록 기준일이 다르지만 KT와 하나로텔레콤으로부터 제공받은 유선전화번호 데이터베이스를 편집하여 발행된다는 점에 착안하여, 특정 기준일(예를 들어 2005년 00월 00일)에서의 이 데이터베이스 상의 가입자들의 비등재율을 파악하여 SUPERPAGE의 포함률을 간접적으로 추정할 수 있다.

데이터베이스 비등재율에 대한 자료는 KTD에서 정기적으로 공개하지 않으나 분석결과에 의하면 2005년 12월 31일 기준으로 전국 비등재율(unlisted rate) P_{unl}의 추정치는 다음 (2)와 같다.

$$KTD \text{ 데이터베이스 비등재율 } P_{unl} \approx 0.33(33\%) \tag{2}$$

이때 추정 결과인 (1)의 $P_{\max} \approx 0.9$와 (2)의 $P_{unl} \approx 0.33$을 식(3)에 대입하여 백분율로 표시되는 SUPERPAGE의 포함률(coverage rate) P_{cov}을 구할 수 있다.

$$\begin{aligned} P_{cov} &= [\{(\text{총가구수} \times P_{\max}) \times (1 - P_{unl})\} / \text{총가구수}] \times 100 \\ &= \{P_{\max} \times (1 - P_{unl})\} \times 100 \\ &= 60.3\% \end{aligned} \tag{3}$$

여기서 총가구수는 통계청(2005)의 인구주택총조사보고서 상의 수치(15,988,274개)이며 총가구수, $P_{\max}$ 및 P_{unl}는 각각 2005년 11월, 12월 및 12월 기준의 비슷한 시점의 수치들이다. $(\text{총가구수} \times P_{\max}) \times (1 - P_{unl})$는 총등재가구수를 나타낸다.

따라서 식(3)은 SUPERPAGE를 표본추출틀로 사용할 때 전체 가구들 중 약 60%만이 조사대상으로 포함된다는 것을 나타낸다.

② 직접적 방법에 의한 포함률 추정

위에서는 인구주택총조사의 총가구수, 정보통신부의 표본조사자료로부터의 $P_{\max}$, 유선전화번호 데이터베이스의 P_{unl}를 이용한 간접적인 방법으로 SUPERPAGE의 포함률을 추정하였다. 이와는 달리 인구주택총조사의 총가구수와 SUPERPAGE 상의 등재가구수를 이용하여 직접적인 방법으로 SUPERPAGE의 포함률을 추정할 수 있다.

SUPERPAGE들은 지역별로 수록 기준일이 서로 상이하므로, 발행된 SUPERPAGE들 중에서 인구주택총조사 기준일(2005년 11월 1일)에 가장 가까운 수록기준일을 갖는 SUPERPAGE들을 이용한다. 이때 백분율로 표시되는 포함률 P_{cov}의 추정식은 식(4)와 같다.

$$P_{cov} = 100 - P_{noncov} \tag{4}$$

여기서 P_{noncov}는 SUPERPAGE의 비포함률(noncoverage rate)이며

$$P_{noncov} = \{(\text{총가구수} - \text{총등재가구수})/\text{총가구수}\} \times 100 \tag{5}$$

로 표현된다. 그리고 총가구수는 통계청(2005)의 인구주택총조사보고서 상의 전국 가구수이며 총등재가구수는 전국의 모든 SUPERPAGE들의 '인명편'에 등재된 전체 가입자수이다.

식(5)에서의 총등재가구수를 구하는 간단한 방법으로서, 각 지역별로 발행되는 SUPERPAGE의 등재가입자수를 식(6)을 이용하여 구한 후 이들을 모두 합하는 방법을 고려할 수 있다.

$$\text{지역별 등재가입자수} = \text{페이지당 성명의 개수} \times \text{총페이지수} \tag{6}$$

그런데 식(6)은 페이지당 기재된 성명의 개수가 동일하거나 거의 동일한 경우 사용할 수 있으나 실제로는 다음과 같은 이유들로 각 페이지당 성명의 수가 차이가 있어 사용할 수 없다.

i) 글자 또는 숫자의 조합인 가입자 주소의 길이가 다르다.

ii) 페이지별 글자(숫자) 크기가 다르다.

iii) 광고지면이 페이지별로 불규칙한 크기로 있다.

vi) 배열을 맞추기 위해 각 행의 폭이 불규칙하게 조정된다.

더욱이 SUPERPAGE의 각 페이지별로 성명의 개수를 수작업으로 직접 확인하는 것은 거의 불가능하므로 식(7)과 같은 상한(upper limits)의 개념을 이용하여 지역별 등재가구수를 추정할 필요가 있다.

$$\text{지역별 등재가입자수의 상한} = \text{기재된 성명의 최대 개수} \times \text{총페이지수} \tag{7}$$

여기서 기재된 성명의 최대 개수는 SUPERPAGE 상의 페이지들 중 성명이 가장 많이 기재된 페이지의 성명들의 개수이다.

그러므로 SUPERPAGE의 비포함률 P_{noncov}을 구하기 위하여 식(5) 대신에 식(7)의 지역별 등재가입자수의 상한을 구한 식(8)에 따라 구한다. 단, 여기서 비포함률은 정확히는 비포함률의 하한(lower limits)을 의미한다.

$$P^{*}_{noncov} = \{(\text{총가구수} - \text{총등재가구수의 상한})/\text{총가구수}\} \times 100 \tag{8}$$

여기서 총등재가구수의 상한은 식(7)의 지역별 등재가입자수의 상한들을 모두 합한 것이며 (총가구수 – 총등재 가구수의 상한)은 비포함 가구수(유선전화 미보유 가구수 + 비등재 가구수)를 나타낸다. 이 비포함 가구수는 정확히는 비포함 가구수의 하한을 의미한다.

따라서 SUPERPAGE의 포함률은 식(4) 대신에 식(9)를 이용하여 구한다.

$$P^{*}_{cov} = 100 - P^{*}_{noncov} \tag{9}$$

전국이 아닌 지역별 SUPERPAGE의 포함률 $P^{*}_{cov,\,area}$을 구하기 위해서는 위의 식(8)과 식(9)를 약간 바꾼 다음 식(10)과 식(11)을 이용한다.

$$P^{*}_{cov,\,area} = 100 - P^{*}_{noncov,\,area} \tag{10}$$

여기서

$$P^{*}_{noncov,\,area} = \{(\text{지역 가구수} - \text{지역 등재가구수의 상한})/\text{지역 가구수}\} \times 100 \tag{11}$$

이고, 지역가구수는 2005년 인구주택총조사보고서 상의 지역 가구수이며, 지역 등재가구수의 상한은 식(7)의 지역 등재가입자수의 상한이다. 만약 지역 가구수가 지역 등재가구수보다 작은 경우 $P^{*}_{noncov,\,area}$은 음수(마이너스) 값이 되므로 $P^{*}_{cov,\,area} > 100$이 된다.

〈표 7〉은 2005년 인구주택총조사 및 SUPERPAGE 전화번호부에 따른 7개 대도시 지역(metropolitan areas)별 가구수, 비포함률 및 포함률의 분포이다. 단, 지역 나열 순서는 총가구수에 따른 것이다.

〈표 7〉에 나타난 바와 같이 대구 지역을 제외한 다른 6개 지역들은 비포함률이 30%를 넘으며 지역별 비포함률 크기의 순서는 지역별 가구수 크기의 순서와 전혀 다르다. 서울은 오히려 비포함률이 인천, 대전, 광주보다 낮으며 이 3개 지역들 중 대전이 가장 높다. 그러므로 포함률은 7개 대도시 지역들 중 대구가 72.4%로 가장 높고 대전이 51.9%로 가장 낮으며 이들 도시 간 약 21%의 차이가 난다.

〈표 8〉은 9개 도 지역(provincial areas)에서의 인구주택총조사 및 SUPERPAGE 전화번호부에 따른 가구수, 비포함률 및 포함률의 분포이다. 단, 도 지역 나열 순서는 총가구수에 따른 것이다.

〈표 8〉과 같이 전라남도와 제주도를 제외한 다른 9개 지역의 비포함률이 20%를 넘고, 도별 비포함률 크기의 순서는 도별 총가구수 크기의 순서와 크게 다르며, 경기도가 가장 높은 40.8%이고 충청북도(35.8%), 강원도(31.9%) 등의 순으로 높다. 이들 지역들 중 비포함률이 가장 높은 경기도와 가장 낮은 제주도 간의 포함률 차이는 23.6%이다.

〈표 7〉 대도시 지역별 가구수, 비포함률 및 포함률 분포

(단위: 호, %)

지역	수록 기준일	가구수 ①	등재가구수 (상한) ②	비포함 가구수 ①-②	비포함률 $P^{*}_{noncov,\,area}$	포함률 $P^{*}_{cov,\,area}$
서울	2006.1.2	3,341,352	2,080,890	1,260,462	37.7	62.3
부산	2005.3.6	1,190,107	803,880	386,227	32.5	67.5
인천	2005.7.20* ~2006.11.10	828,012	490,350	337,662	40.8	59.2
대구	2005.8.31	817,620	591,570	226,050	27.6	72.4
대전	2006.9.5	480,466	249,480	230,986	48.1	51.9
광주	2006.7.12	461,387	262,080	199,307	43.2	56.8
울산	2006.1.14	340,652	219,240	121,412	35.6	64.4

* 인천의 SUPERPAGE는 시 지역 외에 군 지역을 포함하고 있어 전화번호부 수록 기준일이 일정하지 않다.

〈표 8〉 도 지역별 가구수, 비포함률 및 포함률 분포

(단위: 호, %)

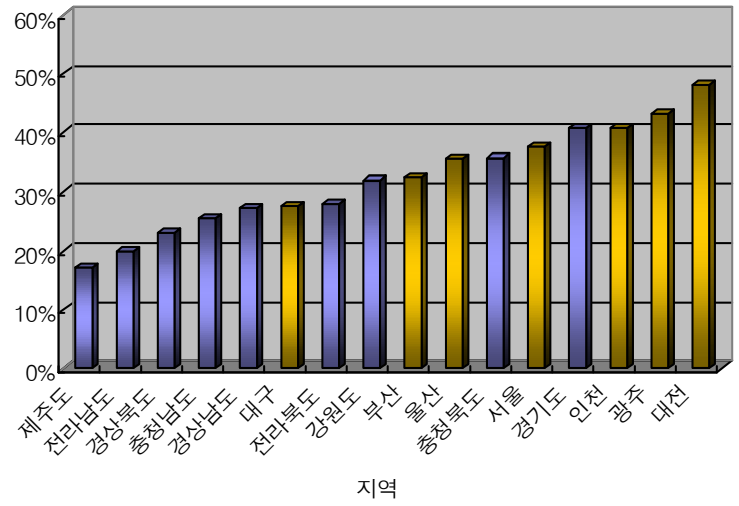
지역	수록 기준일*	가구수 ①	등재가구수 (상한) ②	비포함 가구수 ①-②	비포함률 $P^{*}_{noncov,area}$	포함률 $P^{*}_{cov,area}$
경기도	2004.5.5 ~2006.9.5	3,361,657	1,991,284	1,370,373	40.8	59.2
경상남도	2004.12.11 ~2006.7.7	1,061,259	772,297	288,962	27.2	72.8
경상북도	2004.12.6 ~2005.12.27	943,135	724,711	218,424	23.2	76.8
전라남도	2006.1.1 ~2006.4.19	667,962	535,480	132,482	19.8	80.2
충청남도	2004.12.4 ~2005.12.6	663,445	495,202	168,243	25.4	74.6
전라북도	2006.1.21 ~2006.4.15	621,735	448,119	173,616	27.9	72.1
강원도	2005.5.5 ~2006.5.1	522,225	355,526	166,699	31.9	68.1
충청북도	2005.7.1 ~2005.11.15	507,541	326,034	181,507	35.8	64.2
제주도	2006.8.30 ~2006.9.6	179,719	148,862**	30,857	17.2	82.8

* 각 도의 시·군 지역별 전화번호부 수록 기준일이 일정하지 않다.

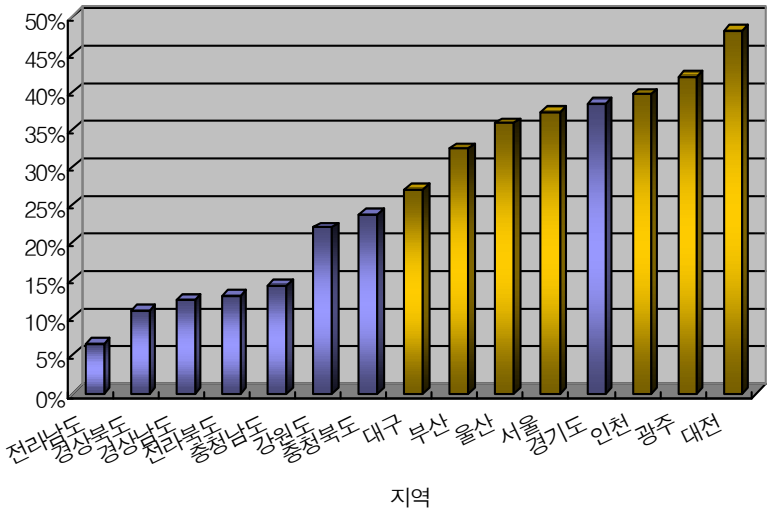
** 제주도 지역 전화번호부는 각 페이지에 가입자명(주택용)과 상호명(업무용)이 같이 수록되어 있어 전체 467개 페이지들 중 약 28%인 132개 페이지를 단순임의추출법(simple random sampling)으로 추출하여 가입자명수(주택용)를 추정한 값이 '148,862'이다.

〈그림 1〉은 〈표 7〉과 〈표 8〉의 대도시 및 도 지역 비포함률을 그래프로 나타낸 것이다. 〈그림 1〉에서 볼 수 있는 바와 같이 도 지역의 비포함률이 대도시 지역에 비해 전반적으로 낮으나 경기도와 충청북도의 경우는 일부 대도시 지역보다 오히려 높다.

〈그림 2〉는 각 대도시의 구 또는 군 지역들 및 제주도를 제외한 각 도의 시 또는 군 지역들의 SUPERPAGE 비포함률들에 대한 중위수(median)를 그래프로 나타낸 것으로서 경기도를 제외한 모든 도 지역의 비포함률 중위수가 대도시 지역보다 낮다. 〈그림 1〉과 비교할 때 대도시 지역의 비포함률의 크기 순위는 변동이 없으나, 도 지역의 경우 일부 변동이 있다.



〈그림 7〉 대도시 및 도 지역별 비포함률



〈그림 2〉 대도시 및 도 지역 내의 구, 시, 군 지역들 비포함률 중위수

〈표 9〉 전국 가구수, 비포함률 및 포함률 분포

(단위: 호, %)

총가구수 ①	등재가구수 (상한) ②	비포함 가구수 ①-②	비포함률 P^*_{noncov}	포함률 P^*_{cov}
15,988,274	10,495,005	5,493,269	34.4	65.6

한편, 각 대도시 또는 각 도 지역 내의 구·시·군 지역들의 비포함률의 분포는 균일하지 않으며 비포함률은 대부분 0%~50% 범위이고 전반적으로 구 또는 시 지역들의 비포함률이 군 지역들보다 높다. 〈부록〉에 있는 그림들은 서울, 부산 및 경기도의 구나 시 또는, 군 지역별 SUPERPAGE의 비포함률을 나타낸 것이다. 한 가지 유의할 것은 일부 구나 군 지역은 비포함률이 음수(마이너스) 값을 갖는다는 점이다. 이 지역들은 SUPERPAGE에 등재된 가입자수가 인구주택총조사 상의 가구수보다 많은 지역으로서 주택용이 아닌 업무용 유선전화번호가 다수 포함된 지역이다. 서울 지역을 예로 들어 살펴보면, 비포함률이 음수 값을 갖는 지역은 중구와 종로구이다. 이들 중 중구는 명동을 포함한 서울의 대표적인 상가 밀집지역으로 2005년 인구주택총조사 상의 가구수는 47,760가구이나 SUPERPAGE 상의 등재가입자수는 약 1.4배인 67,410으로서 식(11)에 이들 수치들을 대입하면 비포함률 $P^*_{noncov,\,area}$ 은 약 −41.1%이고, 식(10)의 포함률 $P^*_{cov,\,area}$ 은 100%를 훨씬 초과하는 141.1%가 된다. 따라서 서울 중구 지역의 가구 표본을 추출할 때 SUPERPAGE를 이용하는 경우 가구번호가 아닌 최소 약 41%의 업무용 번호들이 포함될 수 있다는 것에 주의해야 한다.

〈표 9〉는 대도시 지역과 도 지역을 모두 포함한 전국의 가구수, 비포함률 및 포함률이다.

여기서 포함률 65.6%는 SUPERPAGE를 표본추출틀로 사용할 때 전국 가구들 중 약 66%만이 표본으로 추출된다는 것을 나타낸다.

(3) SUPERPAGE의 실제 포함률

전국 규모의 조사에 대해서 위에서 논의된 2가지 방법에 의한 SUPERPAGE의 비포함률과 포함률 추정결과를 요약하면 〈표 10〉과 같다.

〈표 10〉의 2가지 방법에 따른 SUPERPAGE의 포함률의 차이는 추정방법에 기인한 것이다. 그리고 직접적 방법에 의한 포함률 65.6%와 간접적 방법에 의한 포함률 60.3%는 큰 차이가 나지 않으므로 직접적 방법에 의한 추정이 비교적 정확하다고 할 수 있다.

〈표 10〉 2가지 방법에 의한 SUPERPAGE 포함률 추정결과 요약

간접적 방법에 의한 추정		직접적 방법에 의한 추정
비포함(미보유)률 10%	비포함률 39.7%	비포함률 34.4%
비포함(비등재)률 29.7%		
포함률 60.3%		포함률 65.6%

주: 점선 위의 화살표 방향은 비포함률이 실제로는 더 크다는 것을 의미함

그리고 주목해야 할 것은 SUPERPAGE의 실제 비포함률은 다음과 같은 이유들로 〈표 10〉의 점선까지 범위의 비포함률보다 화살표 방향으로 더 커질 수밖에 없고 결과적으로 포함률은 더 줄어든다는 사실이다.

i) 등재된 번호들 중 일부는 주택용(가구용)이 아닌 업무용으로 사용된다 : 간접적 방법과 직접적 방법 모두 해당.
ii) 등재된 번호들 중 일부는 사용하지 않는 번호이다 : 간접적 방법과 직접적 방법 모두 해당.
iii) 일부 가구들은 2개 이상의 전화를 보유하고 있다(참고로, 정보화실태조사보 고서(2002)에 의하면 이러한 가구들의 비율이 3.3%임) : 간접적 방법과 직접적 방법 모두 해당.
iv) 지역별 SUPERPAGE의 등재가입자수의 상한을 이용해 비포함률을 추정하였다 : 직접적 방법 해당.

사실 업무용 유선전화번호는 앞에서 설명한 서울의 중구와 종로구 등과 같이 전국 대부분 지역의 SUPERPAGE에 포함될 수 있으며 이러한 번호들로 인하여 SUPERPAGE의 비포함률이 〈표 10〉에 나타난 것보다 낮을 것이다. 이 업무용 전화번호는 전화서비스 가입자들이 가입 시에는 주택용으로 가입을 하고 실제로는 업무용으로 사용하기 때문에 발생한다. 따라서 전국 가구들을 대상으로 한 조사에서 SUPERPAGE의 실제 포함률은 직접적 방법에 의해 추정된 65.6%보다 다소 낮을 것이다. 그리고 강현철 외(2008)는 포함률이 59.3% 이하인 것으로 추정하였는데 이는 본 연구의 추정결과를 간접적으로 뒷받침한다고 할 수 있다. 다만 강현철 외(2008)의 논문에서 '59.3%'는 2007년 KOIS의 유선전화번호 데이터베이스 상에 등재된 번호의 개수를 2005년 인구주택총조사보고서 상의 총가구수로

단순히 나눈 수치로서 KOIS 데이터베이스와 인구주택총조사의 기준 시점은 약 2년의 차이가 있다는 점에 주의할 필요가 있다.

또한, i)과 ii)와 관련한 업무용 또는 사용하지 않는 번호의 비율에 관한 정보는 KT나 KTD 등에서 보유하지 않고 있어 추가적인 연구가 필요하다.

3. 비포함 문제 해결을 위한 외국의 연구

Brick et al.(1995)에 의하면 미국의 포함률은 약 70%였으며 Collins(1999)가 언급한 영국의 포함률은 64% 이하였다. 이들 수치는 위에서 제시한 2005년 국내 SUPERPAGE의 실제 포함률 65.6% 이하보다 다소 높거나 유사하다.

유선전화 비등재 가구로 인한 표본추출틀의 비포함 문제를 해결하기 위해 영국보다는 주로 미국에서 1960년대부터 많은 연구가 이루어졌는데, 주요 연구들을 살펴보면 다음과 같다.

미국에서는 비등재로 인한 비포함 문제를 개선하기 위하여 다양한 표본추출방법(sampling techniques)들이 연구되었다. 그 중에서 임의번호추출법은 가장 대표적인 방법으로서 약 반세기에 걸치는 동안 많은 변천을 해왔다. Cooper(1964)는 비포함 문제를 해결하기 위하여 등재된 지역번호 및 국번(prefixes) 뒤의 4자리 수들을 임의로 추출하는 방법을 처음으로 제안하였는데 이것은 임의번호추출법의 초기 방식으로 알려져 있다. 이 방법의 문제점은 추출된 번호들 중 75~80%가 가구용(주택용) 번호가 아닐 수 있어 실제로 사용하는 데 있어 많은 시간과 비용 그리고, 노력이 든다는 것이었다. 이어서 Mitofsky(1970)와 Waksberg(1978)는 전화번호그룹(groups of phone numbers)을 이용하는 2단계 임의번호추출법(two stage RDD sampling method)을 제안하였고, Potthoff(1987)는 이 방법을 일반화한 방법을 제시하였다. 그리고 Lepkowski & Groves(1986)는 각 전화번호 그룹별 등재번호의 개수에 비례하는 방법으로 표본을 추출하는 방법을 제안하였으며, Casady & Lepkowski(1993)는 층화표본추출방법을 개발했다. Lepkowski & Groves(1986) 그리고 Casady & Lepkowski(1993)이 제시한 방법들은 대표적인 목록 이용 임의번호추출법(list-assisted RDD methods)으로서 2단계 임의번호추출법 방법에 비해 사용이 편리하다. 또한 Brick et al.(1995)는 Casady & Lepkowski(1993)의 방법을 좀더 체계화시켜 전국적으로 3~4%의 가구들만이 제외되는 새로운 목록 이용 임의번호추출법를 개발하여 임의번호추출법이 널리 사용될 수 있는 계기를 마련하였다.

Ⅲ. 결론 및 향후 연구방향

위에서 살펴본 바와 같이, 2005년 말을 기준으로 한 대도시 지역 및 도 지역의 SUPERPAGE 포함률은 약 50~80%이며, 전국적으로는 65.6% 수준으로 추정되나 실제 포함률은 업무용 전화번호 등으로 인하여 이보다 다소 낮을 것이다. 또한 SUPERPAGE 포함률은 지역별로 차이가 있으며 전반적으로 대도시 지역이 도 지역보다 포함률이 낮으나 일부 대도시 지역은 도 지역보다 높다.

위에서 다룬 것처럼 미국의 경우 이미 1990년대 중반에 전국적으로 3~4%의 가구들만이 제외되도록 비포함률을 최소화하는 새로운 임의번호추출법을 개발하여 실사에 사용하고 있는 반면, 국내에서 수행되고 있는 대부분의 전화조사들은 이보다 매우 높은 비포함률을 갖는 SUPERPAGE 또는 SUPERPAGE와 거의 대등한 전화번호 데이터베이스를 표본추출틀로 사용하고 있다. 따라서 서론에서 언급한 바와 같이 비포함률이 20%보다 훨씬 낮아야하는 사회조사 또는 공공정책 관련 조사 그리고 선거여론조사에서 심각한 포함편향으로 인하여 조사의 정확성이 크게 떨어질 수 있으며, 2005년 이후에도 유선전화 미보유 가구들과 비등재 가구들이 계속 증가하고 있다는 점을 고려할 때 현재 표본추출틀의 비포함 문제는 더욱 커질 수 있다.

최근 강현철 외(2008)가 임의번호추출법을 국내 전국 규모의 조사에 처음 도입하였다. 그의 연구는 Cooper(1964)가 처음으로 제안한 임의번호추출법을 사용한 것으로 표본추출틀의 비포함 문제를 해결하고자 진행된 국내 최초의 연구라는 데 의의가 있다. 그러므로 이 연구를 바탕으로 보다 많은 연구들이 진행되어 국내에서도 미국의 경우처럼 대부분의 전화조사연구에서 임의번호추출법의 사용이 보편화될 수 있도록 효율적인 임의번호추출법이 개발되어야 할 것이다.

그리고 이동전화만 보유한 가구들로 인한 비포함 문제도 함께 연구되어져야 한다. 이들 가구들은 유선전화를 보유하지 않으므로 임의번호추출법을 사용하더라도 조사대상에서 제외된다. 위에서 다룬 바와 같이 이 가구들의 비율들은 점차 증가하여 2005년에 약 9%로 추정되었으나 2005년 이후의 비율이나 구체적인 특성은 전혀 연구된 바가 없다. 조사의 정확성을 높이기 위하여 이들 가구들이 조사대상에 포함될 수 있도록 Lepkowski & Kim (2005)이 제시한 바와 같이 유선전화 보유자와 이동전화 보유자를 함께 조사하는 방법도 모색되어야 한다.

아울러 인터넷전화(Voice over Internet Protocol: VOIP)에 대한 연구자들의 관심이 필요하다. 김민정(2007)에 의하면 2006년에 일본의 인터넷전화 가입자수가 약 1,500만 명

이었으며, 국내는 2007년에 가입자수가 약 88만 명으로 기존 유선전화의 대체수단으로서 사용이 점차 크게 증가할 것으로 예상된다. 따라서 인터넷전화의 경우도 이동전화와 마찬가지로 비포함 문제에 영향을 줄 수 있으며, 앞으로 이에 대한 연구가 병행되어야 할 것이다.

참고문헌

강현철·한상태·김지연·정용찬·허명회. 2008. "RDD 전화조사와 주요결과." 《조사연구》 9(1): 1-22.

김민정. 2007. "모바일 VoIP 진화와 시사점: 이동전화 시장을 중심으로." 《정보통신정책》 18(4): 18-40.

김선웅. 2004. "이동전화 확산에 따른 유선전화 가구보유율의 변화: 한국을 포함한 주요 국가들을 중심으로." 《조사연구》 5(1): 27-49.

정보통신부. 2005. 《2005년 하반기 정보화실태조사보고서》.

정보통신부. 2006. 《2006년 상반기 정보화실태조사보고서》.

통계청. 2000. 《사회통계조사보고서》.

통계청. 2001. 《정보화실태조사보고서》.

통계청. 2002. 《정보화실태조사보고서》.

통계청. 2005. 《인구주택총조사보고서》.

허명회·강용수·손은진. 2003. "사회조사에서의 가중 칸 설정에 관한 연구." 《한국조사연구학회 추계 학술발표논문집》 57-69.

허명회·강용수·손은진. 2004. "사회조사에서 조사방법에 따른 가중 칸 설정에 관한 연구 : 전화조사, 인터넷 조사, 모바일 조사." 《조사연구》 5(1): 1-26.

Blankenship, A. B. 1977. "Listed versus Unlisted Numbers in Telephone Survey Samples." *Journal of Advertising Research* 17(1): 39-42.

Blumberg, S. J., Luke, J. V., Cynamon, M. L. and Frankel, M. R. 2008. "Recent Trend in Household Telephone Coverage in the United States." *Telephone Survey Methodology* 56-86 (Lepkowski, J. M. et al., eds.). New York: John Wiley and Sons.

Brick, J. M., Waksberg, J., Kulp, D. and Starer, A. 1995. "Bias in List-Assisted Telephone Samples." *American Association of Public Opinion Research* 59(2): 218-235.

Casady, R. J. and Lepkowski, J. M. 1993. "Stratified Telephone Survey Designs." *Survey Methodology* 19: 103-113.

Collins, M. 1999. "Editorial: Sampling for UK Telephone Surveys." *Journal of the Royal Statistical Society Series A* 162: 1-4.

Collins, M., and Sykes, W. 1987. "The Problems of Non-Coverage and Unlisted Numbers in Telephone Surveys in Britain." *Journal of the Royal Statistical Society Series A* 150(3):

241-253.

Cooper, S. L. 1964. "Random Sampling by Telephone: An Improved Method." *Journal of Marketing Research* 1(4): 45-48.

Kim, Sun Woong and Lepkowski, J. M. 2002a. "Telephone Household Non-Coverage and Mobile Telephones." presented at the Annual Meeting of the American Association for Public Opinion Research, St. Pete, FL.

Kim, Sun Woong and Lepkowski, J. M. 2002b. "The Changes and Trends in Telephone Household Coverage." *Proceedings of the Section on Survey Research Methods.* American Statistical Association 1851-1856.

Kuusela, V. 2003. "Mobile Phones and Telephone Survey Methods." *Proceedings of the Fourth Association of Survey Computing International Conference* 317-327.

Kuusela, V., Callegaro, M. and Vehovar, V. 2008. "The Influence of Mobile Telephones on Telephone Surveys." *Telephone Survey Methodology* 87-112(Lepkowski, J. M., et al., eds.). New York: John Wiley and Sons.

Lepkowski, J. M. and Groves, R. M. 1986. "A Two Phase Probability Proportional to Size Design for Telephone Sampling." *Proceeding of the Section on Survey Research Methods.* American Statistical Association 73-98.

Lepkowski, J. M. 1988. "Telephone Sampling Methods in the United States." *Telephone Survey Methodology* 73-98(Groves, R. M. et al., eds.). New York: John Wiley and Sons.

Lepkowski, J. M., Tucker, C., Michael, B., de Leeuw, E. D., Japec, L., Lavrakas, P. J., Link, M. W. and Sangster, R. L. 2008. *Telephone Survey Methodology.* New York: John Wiley and Sons.

Lepkowski, J. M. and Kim, Sun Woong 2005. "Dual-frame Landline/Cellular Telephone Survey Design." presented in the invited session of the *Section on Survey Research Methods.* American Statistical Association.

Mitofsky, W. 1970. "Sampling of Telephone Households." Unpublished CBS Memorandum.

Moberg, P. E. 1982. "Biases in Unlisted Phone Numbers." *Journal of Advertising Research* 22(4): 51-55.

Office of Telecommunications United Kingdom. 2003. "Consumers' Use of Fixed and Mobile Telephony, Q13 May."
http://www.oftel.gov.uk/publications/research/2003/q13fix&mobr0703.htm.

Potthoff, R. F. 1987. "Some Generalizations of the Mitofsky-Waksberg Technique for Random Digit Dialing." *Journal of the American Statistical Association* 82(398): 409-418.

Thornberry, O. T. and Massey, J. T. 1988. "Trends in United States Telephone Coverage Across Time and Subgroups." *Telephone Survey Methodology* 25-50(Groves, R. M., et al., eds.). New York: John Wiley and Sons.

Trewin, M. and Lee, G. 1988. "International Comparisons of Telephone Coverage." *Telephone Survey Methodology* 9-24(Groves, R. M., et al., eds.). New York: John Wiley and Sons.

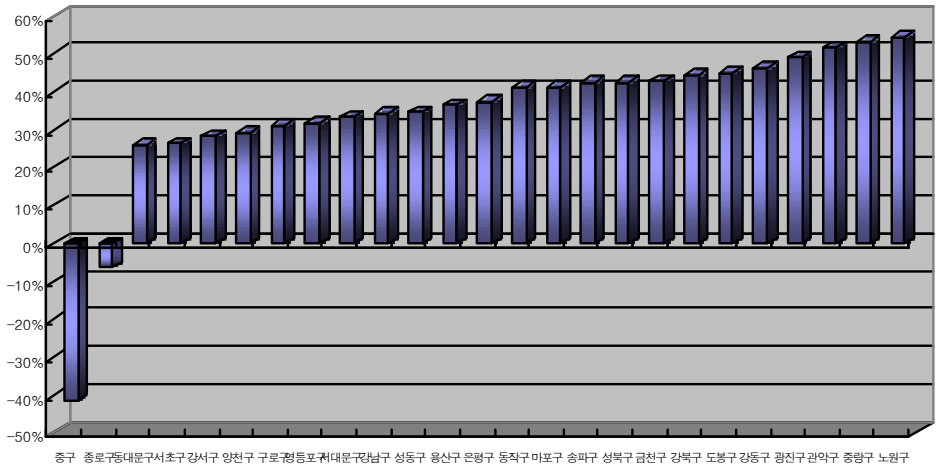
Tucker, C. 2005. "Household Telephone Service and Usage Patterns in the United States in 2004." presented at Cell Phone Summit. New York: New York.

Waksberg, J. 1978. "Sampling Methods for Random Digit Dialing." *Journal of the American Statistical Association* 73: 40-46.

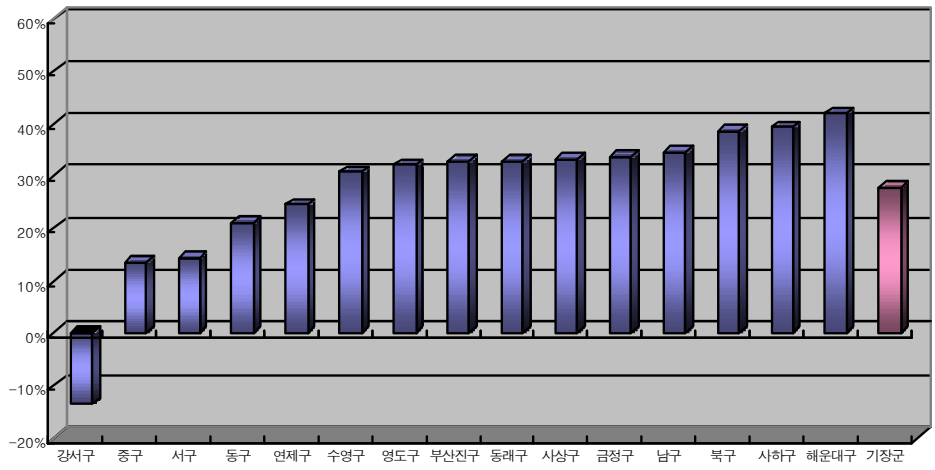
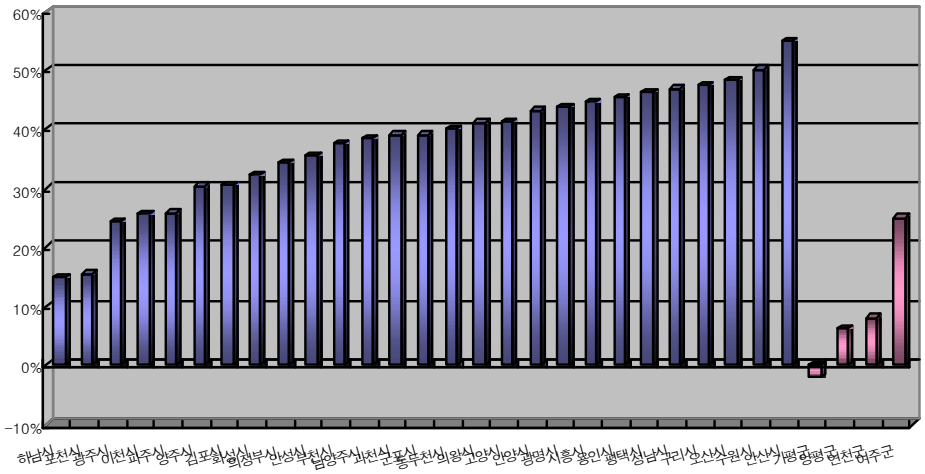
[접수 2008/9/19, 1차수정 2008/11/18, 2차수정 2009/2/2,
게재확정 2009/2/23]

〈부록〉 서울, 부산 및 경기도 지역의 구·시·군별 비포함률

서울



부산



경기도