

연구논문

임의번호걸기와 시간균형할당표집에 의한 전화조사의 주요결과

The Major Findings of the Telephone Survey by
Random Digit Dialing and Time-Balanced Quota Sampling

허명희* · 한상태** · 김지연*** · 성은하**** · 강현철*****

M. H. Huh · S. T. Han · J. Y. Kim · E. H. Sung · H. Kang

최근 우리나라의 전화조사가 전화번호부 기반에서 벗어나 임의번호걸기(RDD)로 이행되는 추세에 있다. 그러나 아직도 대부분의 전화조사가 지역·성·나이대 할당표집으로 수행되므로 이로 인해 재택성향이 큰 사회적 계층에 편중되기 쉽다는 지적이 있어 왔다. 이에 대한 대응으로서 시간균형할당표집이 제안된 바 있었으나 실증적 조사는 이제까지 부재하였다(허명희·황진모 2006). 이 연구는 TV시청환경에 대한 조사를 RDD와 통상적인 할당표집의 결합방식으로 수행한 결과와 RDD와 시간균형할당표집의 결합방식으로 수행한 결과를 비교하여 제시한다.

주제어: 전화조사, 임의번호걸기(RDD), 시간균형할당표집

Korean telephone surveys have been based on telephone directory and thus criticized for considerable under-coverage. Now, Korean survey institutions progress to random digit dialing (RDD) very actively. But still most surveys are administered by quota sampling, prone to assign heavier weights to social classes with more hours staying indoor. As a practical remedy, time-balanced quota sampling scheme was proposed by Huh and Hwang (2006). This study compares two telephone surveys on TV audience environment in Korea: RDD with conventional quota sampling versus

* 고려대학교 통계학과 교수

** 호서대학교 정보통계학과 교수

*** (주)밀워드브라운미디어리서치 상무

**** 호서대학교 정보통계학과 석사

***** 교신저자(corresponding author): 호서대학교 정보통계학과 부교수 강현철.

E-mail: hychkang@hoseo.edu

RDD with time-balanced quota sampling.

key words: telephone survey, random digit dialing(RDD), time-balanced quota sampling

I. 서론

우리나라 전화조사는 1980년대 후반 이래 선거여론조사를 비롯한 각종 사회조사에서 주된 방법으로서의 역할을 해 내고 있다. 수년 전부터 조사전문기관들이 효율적인 실사를 위해 CATI(computer-assisted telephone interview) 시스템을 도입함으로써 실사의 질이 한층 높아졌고 조사의 신뢰도에 대한 사회적 평가도 양호한 편이다. 그럼에도 불구하고 우리나라 전화조사는 몇 가지 문제점을 가지고 있다.

첫째, 국내의 전화조사는 모두 전화번호부 또는 이것의 디지털 버전인 CD 전화번호부를 활용하여 이루어지는데, 가구 기준으로 우리나라 전화번호부의 포함률(coverage)은 60% 미만으로 추정된다(허명희 외 2007; 강현철 외 2008). 이처럼 낮은 포함률은 응답자 선택편향을 초래하여 대표성 확보에 상당한 문제를 유발할 수 있다. 전화번호부의 포함률이 낮은 원인은 비등재(unlisting)와 비보유(no telephone) 가구 때문이다. 사생활 침해의 우려 및 개인정보 보호에 대한 민감성 증대로 인해 비등재 비율의 증가가 예상되는 외에 1인 가구의 유선전화 비보유 추세가 확대일로에 있으므로 포함률은 더욱 낮아질 것으로 예상된다(허명희 2007).

전화번호부 표본에 대한 대안 중 하나는 임의번호걸기(RDD: random digit dialing) 조사이다. 이 방법은 전화번호를 가입자 리스트에서가 아니라 컴퓨터가 임의추출한다. 강현철 외(2008)의 연구에서는 임의번호걸기 방법을 이용해 전화조사를 실시하였는데, 주요 결과는 다음과 같았다. 1) 전화번호부의 모집단 포함률은 60% 이하로 추정되었다. 2) 전화번호부 비등재 가구가 등재 가구에 비해 소득이 다소 많은 것으로 나타났다. 3) 비등재 번호 응답자들이 등재 번호 응답자들에 비해 개인정보에 대한 침해 우려 수준이 다소 높았으며, 4) 이념성향에서 진보적인 경향이 있었다. 이와 같은 사실로부터 전화번호부에 기초한 표본추출이 사회경제적 측면에서 편향된 결과를 산출해 내고 있음을 알 수 있다.

둘째, 할당표집(quota sampling)의 문제이다. 우리나라 전화조사는 할당표집에 의해

지역·성·나이대의 각 조합에 모집단 크기에 비례하는 표본을 배분한다. 조사를 신속하게 진행할 수 있다는 점 때문에 대부분의 조사전문기관들이 이 방법을 채택하고 있지만 다음과 같은 문제가 내포되어 있다(허명희·황진모 2006). 1) 지역·성·나이대를 제어한다고 하더라도 낮 시간 채택자 중심의 조사자료가 만들어진다는 점이다. 낮 시간 채택 성향과 직업·교육수준 등 주요 사회경제적 속성이 연관되므로 조사자료가 편향될 가능성이 있다. 2) 조사가 진행됨에 따라 점차 조사 달성률이 떨어지게 된다는 점이다. 마감 시각을 앞 둔 2~3 시간 전에는 빈 칸을 채우는 데 급급하기 때문이다.

이러한 채택성향의 불균형을 개선하기 위하여 허명희·황진모(2006)는 시간균형할당 표집(time-balanced quota sampling)을 제안하였는데, 주요 결과를 살펴보면 다음과 같다. 통계청의 2004년 ‘생활시간조사’ 자료를 토대로 기존의 지역·성·나이대 조합에 조사 시간대 할당을 추가함으로써 응답자 선택편향의 상당 부분이 제거됨을 보였다. 이 방법은 조사회사에 추가적 부담을 주지 않는다. 단지 낮 시간대 비채택 가구들의 리스트를 만들어야 하므로 CATI 시스템 세팅을 조금 변경하기만 하면 된다.

본 연구에서는 TV시청환경에 대한 조사를 강현철 외(2008)의 임의번호결기와 시간균형할당표집에 의한 전화조사로 실시하여 그의 주요 결과를 임의번호결기와 통상적인 할당표집에 의한 전화조사의 결과와 비교하여 제시한다.

II. 임의번호결기와 시간균형할당표본

본 연구에서는 다음 두 독립적인 조사 표본으로부터 얻은 결과들을 비교할 것이다.

임의번호결기만을 고려한 강현철 외(2008)의 연구에서는 주중과 주말 각각에서 2,500명을 목표 표본크기로 정하고 조사를 실시하였는데, 주중의 조사에서는 월, 화, 수요일 3일 동안(2007.8.20~22) 2,392명, 주말의 조사에서는 토요일과 일요일 동안(2007.8.18~19) 2,466명에 대해 조사를 성공적으로 완료하였다. 이하 이 조사를 ‘RDD. Basic’으로 지칭하기로 한다.

본 연구를 위하여 임의번호결기와 시간균형할당을 동시에 고려하여 조사를 실시하였는데, 1,000명을 목표 표본크기로 정하고 조사를 실시하였으며, 화요일(2007.10.9) 낮 시간대(오후 3시~6시 30분)와 저녁시간대(6시 30분 이후)로 나누어 조사가 진행되었다. <표 1>에서 제시된 것과 같이, 낮 시간대와 저녁 시간대의 목표 표본크기는 각각 358명,

642명인데 이는 채택율을 고려하여 얻어진 수치이다. 최종적으로 855명에 대하여 조사를 성공적으로 완료하였다. 이하 이 조사를 'RDD,Balanced'로 지칭하기로 한다.

〈표 1〉 조사 시간대별 표본할당

(단위: 명)

| 조사 시간대 | 구 분 | 합 계 | 20대 | 30대 | 40대 | 50대 | 60대 |
|----------------------|-----|-------|-----|-----|-----|-----|-----|
| 낮 시간대 (6시 30분 이전) | 남 | 133 | 37 | 24 | 26 | 24 | 22 |
| | 여 | 225 | 40 | 62 | 53 | 36 | 34 |
| | 계 | 358 | 77 | 86 | 79 | 60 | 56 |
| 저녁시간대 (6시 30분 이후) | 남 | 371 | 83 | 105 | 99 | 56 | 28 |
| | 여 | 271 | 74 | 62 | 67 | 44 | 24 |
| | 계 | 642 | 157 | 167 | 166 | 100 | 52 |
| 계 | 남 | 504 | 120 | 129 | 125 | 80 | 50 |
| | 여 | 496 | 114 | 124 | 120 | 80 | 58 |
| | 계 | 1,000 | 234 | 253 | 245 | 160 | 108 |

〈표 2〉 유형별 전화번호의 수

(단위: 명)

| 유 형 | RDD.Basic | | | | RDD.Balanced | | |
|----------|---------------|-------------------|-------------------|-------------------|-------------------|------------------|-------------------|
| | 주말 조사 | | 주중 조사 | | | | |
| 유효 번호 | 부재중 | 14,602 (24.8%) | 28,744 (48.9%) | 16,988 (23.8%) | 34,892 (49.0%) | 5,431 (24.2%) | 11,589 (51.7%) |
| | 통화중 | 1,561 (2.7%) | | 3,379 (4.7%) | | 1,450 (6.5%) | |
| | 조사거부/ 중단 | 7,001 (11.9%) | | 7,718 (10.8%) | | 2,391 (10.7%) | |
| | 쿼터불일치 | 3,114 (5.3%) | | 4,415 (6.2%) | | 1,462 (6.5%) | |
| | 조사완료 | 2,466 (4.2%) | | 2,392 (3.4%) | | 855 (3.8%) | |
| 무효 번호 | 결 번 | 26,811 (45.6%) | 30,026 (51.1%) | 31,168 (43.7%) | 36,385 (51.0%) | 9,507 (42.4%) | 10,807 (48.3%) |
| | 상호번호 | 1,745 (3.0%) | | 3,273 (4.6%) | | 710 (3.2%) | |
| | FAX 등 기타번호 | 1,470 (2.5%) | | 1,944 (2.7%) | | 590 (2.6%) | |
| 계 | 58,770 | | 71,277 | | 22,396 | | |

조사를 완료하기 위해 통화를 시도한 전화번호의 개수, 유효번호, 무효번호의 수는 <표 2>와 같으며, 두 방법에서 전화번호의 유형별 비율은 큰 차이가 없는 것으로 나타났다. 두 방법 모두 전화번호부 기반의 조사에 비해 대략 2배 정도의 통화시도 횟수가 필요할 것으로 추정된다. 통화시도 중 유효번호의 비율이 대략 50%이기 때문이다. 그러나 CATI 시스템이 유효/무효번호의 여부를 파악하는 데 걸리는 시간은 몇 초에 불과하므로, 전체 조사시간에 미치는 영향은 크지 않다고 할 수 있다. 또한 시간대별로 할당된 표본크기가 있으므로 시간균형할당의 RDD 조사에서는 쿼터불일치의 비율이 높을 것이라 예상했었으나, RDD 조사의 쿼터불일치 비율과 큰 차이가 없었다.

<표 3> 목표 표본크기 대비 조사 달성률

(단위: 명)

| | | RDD.Basic | | | | RDD.Balanced | |
|-----|----------|-----------|--------|-------|--------|--------------|--------|
| | | 주말 조사 | | 주중 조사 | | | |
| 전 체 | | 2,466 | 98.7% | 2,392 | 95.7% | 855 | 85.5% |
| 성 | 남 자 | 1,195 | 96.8% | 1,129 | 91.5% | 382 | 75.8% |
| | 여 자 | 1,271 | 100.4% | 1,263 | 99.8% | 473 | 95.4% |
| 연 령 | 19~29세 | 528 | 97.4% | 485 | 89.5% | 140 | 59.8% |
| | 30대 | 570 | 97.4% | 542 | 92.7% | 219 | 86.6% |
| | 40대 | 554 | 97.9% | 554 | 97.9% | 218 | 89.0% |
| | 50대 | 375 | 100.9% | 365 | 98.1% | 162 | 101.3% |
| | 60세 이상 | 439 | 100.9% | 446 | 102.5% | 116 | 107.4% |
| 지 역 | 서 울 | 516 | 96.4% | 502 | 93.8% | 169 | 77.2% |
| | 인천/경기 | 672 | 100.0% | 649 | 96.7% | 256 | 93.8% |
| | 대전/충남/충북 | 252 | 100.4% | 248 | 98.8% | 83 | 83.8% |
| | 광주/전남/전북 | 261 | 98.1% | 242 | 91.0% | 85 | 83.3% |
| | 대구/경북 | 263 | 98.5% | 254 | 95.1% | 94 | 89.5% |
| | 부산/울산/경남 | 401 | 99.3% | 400 | 99.0% | 139 | 85.8% |
| | 강원/제주 | 101 | 95.3% | 97 | 91.5% | 29 | 72.5% |

목표 표본크기 대비 조사 달성률은 RDD.Basic.주말이 $2,466/2,500 = 98.7\%$ 이고 RDD.Basic.주중이 $2,392/2,500 = 95.7\%$ 이며, RDD.Balanced는 $855/1,000 = 85.5\%$ 이다(〈표 3〉 참조). 이와 같이 목표에 미달하는 응답자가 조사된 이유는 조사시간의 마지막까지 할당표본추출의 원칙을 고수하였기 때문이다. 특히 두 방법 모두에서 성별에서는 남자 그룹, 연령대에서는 낮은 연령대 그룹, 지역적으로는 서울과 강원/제주 지역에서 상대적으로 달성률이 낮았다. 이러한 현상은 실제 채택률과 밀접한 관련이 있는 것으로 생각된다.

RDD.Balanced의 경우 6시 30분 이전의 조사 달성률은 80.4%, 6시 30분 이후의 조사 달성률은 88.3%로 저녁시간대에서 상대적으로 달성률이 높았다. 또한 비교적 평이한 조사 주제였고 충분한 수의 면접원을 투입하였음에도 불구하고 조사의 목표크기를 다 채우지 못하였으며, 대부분의 그룹에서 RDD.Basic 방법보다 낮은 조사 달성률을 보였다. 실제로 시간균형할당으로 인해 낮 시간대에 조사를 더 할 수 있는 고연령층을 저녁시간대에 조사를 진행함으로써 RDD.Basic 방법보다 조사시간이 더 길어지게 되었다. 따라서 시간균형할당 방법을 적용할 경우 조사시작 시간과 투입 면접원수를 좀더 고려해야 할 것으로 판단되며, CATI 시스템의 개선 등을 통해 조사시간을 줄이기 위한 고민과 노력이 필요하다고 생각된다.

RDD.Balanced 표본에서 유효 전화번호(조사완료, 조사거부/중단, 쿼터 불일치, 통화중, 부재중) 11,589개 중 5,018개(43.3%)는 인명편 전화번호에 등재되어 있었고, 6,571개(56.7%)는 등재되어 있지 않았다. 또한 조사완료 전화번호 855개 중 464개(54.3%)는 인명편 전화번호에 등재되어 있었고, 391개(45.7%)는 등재되어 있지 않았다. 특히 조사완료 전화번호 중 서울, 경기, 인천 등 수도권의 등재율은 46.8%인 반면 비수도권의 등재율은 61.6%로 나타났다. 이와 같은 유효번호/조사완료번호 대비 등재율은 강현철 외(2008)의 RDD.Basic 표본에서 나타난 결과와 유사하다.

Ⅲ. 심리적 성향과 사회인구적 요인의 비교

이 절에서는 RDD.Balanced 표본의 응답자에서 사회인구적 특성과 심리적 성향이 어떻게 나타나고 있는지를 RDD.Basic 표본과 비교하여 살펴보고자 한다. 여기서는 공정한 비교를 위해 RDD.Basic 표본의 주중 조사 응답자만을 이용하기로 한다. 또한 조사시

점의 16개 시·도 성/연령별 인구 분포에 맞도록 가중치를 부여하여 결과를 산출하였다.

먼저 4개 심리적 성향에 대한 응답 분포가 <표 4>에 제시되어 있다. 본 연구에서는 개인정보 침해수준에 대한 인식, 삶에 대한 전반적 만족도, 여론조사 결과의 신뢰도, 주관적 이념성향 등 4개의 요인을 살펴보았는데, 4개의 요인 모두에 대하여 응답 분포가 통계적으로 유의한 차이를 보이지 않았다.

<표 4> 심리적 성향에 대한 응답 분포

(단위: 명)

| 요 인 | 범 주 | RDD.Basic | | RDD.Balanced | | χ^2 |
|-------------------------|--------|-----------|-------|--------------|-------|--|
| 개인정보 침해수준에 대한 인식* | 매우 심각 | 792 | 37.2% | 323 | 40.3% | $\chi^2 = 4.195$ $p\text{-value} = 0.380$ |
| | 심 각 | 927 | 43.5% | 320 | 39.8% | |
| | 보 통 | 342 | 16.0% | 128 | 16.0% | |
| | 불심각 | 62 | 2.9% | 29 | 3.6% | |
| | 매우 불심각 | 8 | 0.4% | 3 | 0.3% | |
| 삶에 대한 전반적 만족도* | 매우 만족 | 122 | 5.2% | 54 | 6.4% | $\chi^2 = 1.653$ $p\text{-value} = 0.799$ |
| | 만 족 | 818 | 34.7% | 290 | 33.9% | |
| | 보 통 | 1,050 | 44.5% | 377 | 44.2% | |
| | 불만족 | 333 | 14.1% | 119 | 1.5% | |
| | 매우 불만족 | 37 | 1.6% | 13 | 14.0% | |
| 여론조사 결과의 신뢰도* | 매우 신뢰 | 29 | 1.3% | 12 | 1.4% | $\chi^2 = 5.084$ $p\text{-value} = 0.279$ |
| | 신뢰 | 491 | 22.0% | 158 | 19.1% | |
| | 보 통 | 1,156 | 51.8% | 423 | 51.2% | |
| | 불신뢰 | 474 | 21.2% | 196 | 23.8% | |
| | 매우 불신뢰 | 83 | 3.7% | 37 | 4.5% | |
| 주관적 이념성향* | 매우 보수적 | 57 | 2.7% | 21 | 2.7% | $\chi^2 = 1.221$ $p\text{-value} = 0.875$ |
| | 보수적 | 638 | 30.3% | 227 | 29.0% | |
| | 중 간 | 963 | 45.7% | 355 | 45.3% | |
| | 진보적 | 422 | 20.0% | 168 | 21.4% | |
| | 매우 진보적 | 29 | 1.4% | 13 | 1.6% | |

* 무응답을 제외하고 분석되었음.

〈표 5〉 사회인구적 요인에 대한 응답 분포

(단위: 명)

| 요 인 | 범 주 | RDD.Basic | | RDD.Balanced | | χ^2 |
|------------------|-----------|-----------|-------|--------------|-------|---|
| 가구원수 | 1 | 135 | 5.6% | 52 | 6.0% | $\chi^2 = 5.631$ $p\text{-value} = 0.344$ |
| | 2 | 386 | 16.1% | 110 | 12.8% | |
| | 3 | 472 | 19.7% | 178 | 20.8% | |
| | 4 | 978 | 40.9% | 357 | 41.8% | |
| | 5 | 322 | 13.5% | 119 | 13.9% | |
| | 6명 이상 | 99 | 4.1% | 40 | 4.7% | |
| 19세 이상 가구원수 | 1 | 168 | 7.0% | 66 | 7.7% | $\chi^2 = 1.963$ $p\text{-value} = 0.854$ |
| | 2 | 1,157 | 48.3% | 393 | 46.0% | |
| | 3 | 513 | 21.4% | 192 | 22.4% | |
| | 3 | 421 | 17.6% | 158 | 18.5% | |
| | 5 | 109 | 4.6% | 36 | 4.2% | |
| | 6명 이상 | 26 | 1.1% | 10 | 1.2% | |
| TV 수상기 대수 | 0 | 28 | 1.2% | 6 | 0.7% | $\chi^2 = 4.503$ $p\text{-value} = 0.212$ |
| | 1 | 1,318 | 55.1% | 444 | 51.9% | |
| | 2 | 889 | 37.2% | 341 | 39.9% | |
| | 3대 이상 | 157 | 6.6% | 64 | 7.5% | |
| 케이블 TV 시청 여부* | 예 | 1,665 | 71.1% | 616 | 72.8% | $\chi^2 = 0.872$ $p\text{-value} = 0.350$ |
| | 아니오 | 676 | 28.9% | 230 | 27.2% | |
| 위성 TV 시청 여부* | 예 | 280 | 12.1% | 107 | 12.7% | $\chi^2 = 0.236$ $p\text{-value} = 0.627$ |
| | 아니오 | 2,043 | 87.9% | 736 | 87.3% | |
| 가구 월소득* | 100만원 이하 | 370 | 19.9% | 89 | 12.2% | $\chi^2 = 31.184$ $p\text{-value} = 0.001$ |
| | 101~200만원 | 332 | 17.8% | 119 | 16.3% | |
| | 201~300만원 | 482 | 25.9% | 205 | 28.2% | |
| | 301~400만원 | 283 | 15.2% | 151 | 20.7% | |
| | 401~500만원 | 173 | 9.3% | 83 | 11.4% | |
| | 501만원 이상 | 222 | 11.9% | 81 | 11.1% | |
| 응답자 학력* | 중졸 이하 | 493 | 21.1% | 124 | 14.9% | $\chi^2 = 17.496$ $p\text{-value} = 0.001$ |
| | 고 졸 | 744 | 31.9% | 285 | 34.2% | |
| | 대졸/대재 | 987 | 42.3% | 393 | 47.1% | |
| | 대학원 이상 | 108 | 4.6% | 32 | 3.8% | |

* 무응답을 제외하고 분석되었음.

다음으로 7개 사회인구적 특성에 대한 응답 분포가 <표 5>에 제시되어 있다. 본 연구에서 고려한 요인들 중 가구원수, 19세 이상 가구원수, TV 수상기 대수, 케이블 TV 시청 여부, 위성 TV 시청 여부 등에 대하여는 응답 분포가 통계적으로 유의한 차이를 보이지 않았다.

특히 두 표본의 가구원수 분포는 통계적으로 유의한 차이를 보이지 않았으나($\chi^2=5.631$, p -value = 0.344), 두 표본 모두 2005년 인구주택총조사의 결과와는 상당히 다른 분포를 보이고 있다(2005년 인구주택총조사의 가구원수 분포는 1인: 20%, 2인: 22%, 3인: 21%, 4인: 27%, 5인: 8%, 6인 이상: 2%이며, 두 표본의 조사시점별 인구구성과 다를 수 있음). 즉, 가구원수 3인 이하의 가구는 모집단의 분포에 비하여 과소하게 표집되었고, 가구원수 4인 이상의 가구는 모집단의 분포에 비하여 과다하게 표집되었음을 알 수 있다. 이러한 현상은 현재의 집전화 기반 전화조사가 갖는 한계라고 할 수 있는데, 1~2인 가구에 속한 사람들은 상대적으로 젊은 세대이고 집 밖에서 머무는 시간이 많으므로 접촉이 상대적으로 어려우며, 집전화 자체를 보유하지 않은 경우가 많아서 접촉이 불가능한 비율도 상당히 높을 것으로 추측된다.

두 표본의 가구 월소득 분포는 통계적으로 유의한 차이를 보였다($\chi^2=31.184$, p -value = 0.001). 특히 RDD.Balanced의 경우 RDD.Basic에 비하여 200만원 이하의 소득층이 상대적으로 적게, 200만원 이상의 소득층이 상대적으로 많게 표집되었다. 두 표본의 학력 분포도 통계적으로 유의한 차이를 보였다($\chi^2=17.496$, p -value = 0.001). RDD.Balanced의 경우 RDD.Basic에 비하여 중졸 이하의 학력층이 상대적으로 적게, 고졸과 대졸의 학력층이 상대적으로 많게 표집되었다.

한편, RDD.Balanced 표본에 대하여 낮 시간대(오후 6시 30분 이전)와 저녁 시간대(오후 6시 30분 이후) 구분에 따른 사회인구적 특성을 비교하여 보았는데 가구 월소득과 응답자 학력 분포에서 통계적으로 유의한 차이를 보였다(<표 6> 참조). 가구 월소득의 경우 낮 시간대에는 300만원 이하의 소득층이, 저녁 시간대에는 300만원 이상의 소득층이 상대적으로 많게 표집되었다. 응답자 학력의 경우 낮 시간대에는 고졸 이하의 학력층이, 저녁 시간대에는 대졸 이상의 학력층이 상대적으로 많게 표집되었다.

〈표 6〉 RDD.Balanced 표본의 시간대별 응답 분포

(단위: 명)

| 요 인 | 범 주 | 낮 시간대 | | 저녁 시간대 | | χ^2 |
|------------|-----------|-------|-------|--------|-------|---|
| 가구 월소득* | 100만원 이하 | 40 | 17.6% | 50 | 10.0% | $\chi^2=16.535$ $p\text{-value}=0.005$ |
| | 101~200만원 | 38 | 16.7% | 81 | 16.1% | |
| | 201~300만원 | 72 | 31.7% | 133 | 26.5% | |
| | 301~400만원 | 34 | 15.0% | 117 | 23.3% | |
| | 401~500만원 | 20 | 8.8% | 64 | 12.7% | |
| | 501만원 이상 | 23 | 10.1% | 57 | 11.4% | |
| 응답자 학력* | 중졸 이하 | 54 | 20.9% | 71 | 12.3% | $\chi^2=17.085$ $p\text{-value}=0.001$ |
| | 고 졸 | 91 | 35.3% | 194 | 33.7% | |
| | 대졸/대재 | 110 | 42.6% | 283 | 49.1% | |
| | 대학원 이상 | 3 | 1.2% | 28 | 4.9% | |

* 무응답을 제외하고 분석되었음.

그동안 조사업계 및 학계의 여러 연구에서 저소득층, 저학력층이 과다하게 표집됨으로써 전화조사의 결과가 편향될 수 있다는 것이 현행 전화조사의 문제점 중 하나로 지적되어 왔다(허명희 2007; 홍성준 외 2009; 우정엽 외 2011). 강현철 외(2008), 허명희·김영원(2008), 우정엽 외(2011)의 연구는 임의번호걸기를 사용함으로써 이러한 표집 편향을 일부 제어할 수 있음을 보였으며, 본 연구는 시간균형할당을 함께 사용함으로써 표집 편향을 더욱 줄일 수 있다는 것을 확인하여 주었다.

IV. 맺음말

전화조사는 1980년대 후반 이래 우리나라에서 사회여론조사의 주된 방법론으로 활용되고 있으며, 학계와 조사업계의 많은 경험과 연구를 통해 조사기법 또한 많은 발전을 이루었다. 또한 주요 조사전문기관들은 최근 몇 년 사이에 CATI 시설을 갖추어 조사의 효율과 실사의 질을 높여 왔다. 그럼에도 불구하고 우리나라 전화조사는 몇 가지 문제점을 가지고 있다. 특히 전화번호부 DB를 표본추출 틀로 사용하는 것에 대하여 지속적인 문제 제기가 있어 왔고, 이에 대한 대안으로 임의번호걸기에 관심이 집중되고 있으며 최

근 다수의 실증 연구에서 임의번호걸기가 정확성 제고에 효과적임이 확인되었다. 또한 본 연구에서는 임의번호걸기에 시간균형할당표집을 함께 적용함으로써 조사표본의 편향성을 더 줄일 수 있음이 보여졌다.

최근 우리나라 전화조사는 낮은 전화번호부 등재율뿐만 아니라 통신기기의 다양화에 따른 포함률의 저하 및 조사환경의 변화 등 새로운 위협 요인들에 직면하고 있다. 한 가지는 유선전화를 보유하지 않은 가구의 비율이 상당히 높다는 것이다. 특히 가구원이 1~2명인 가구에서는 유선전화를 보유하지 않고 무선전화만을 사용하는 경우가 매우 많다. 다른 한 가지는 집전화 중 070 등으로 시작하는 인터넷 전화의 비율이 급격하게 높아지고 있다는 점인데, 인터넷 전화를 사용하는 가구 역시 젊은 세대일 가능성이 많다. 따라서 현재의 임의번호걸기 표집 틀에도 무시할 수 없는 과소포함(under-coverage) 문제가 있다. 그런데 이 문제는 점점 심각해질 것이며, 집전화 기반의 가구조사에서 가구원수의 편향이 조사결과에 미치는 영향과 가중치 부여 등에 대하여 보다 심도 있는 연구가 필요하다.

조사방식에 대하여 인터넷 전화번호가 포함된 임의번호걸기 표집 틀을 구축하는 방식과 1~2인 가구에 대하여 무선전화조사를 병행하는 방식 등에 대하여 구체적 제안이 필요하다. 또한 이러한 표집 틀에 기반할 경우에도 시간균형할당표집 등 표본배분 방식과 가중치 적용 등에 대한 다수의 실증적 연구가 필요하다.

참고문헌

- 강현철·한상태·김지연·정용찬·허명희. 2008. “RDD 전화조사와 주요결과.” 《조사연구》 9(1): 1-22.
- 우정엽·김지운·문중배. 2011. “표집 틀 설정과 표본추출방법에 따른 정치성향 분석의 문제점: 임의번호걸기(Random Digit Dialing)과 전화번호부 추출방법 비교.” 《조사연구》 12(1): 153-174.
- 허명희. 2007. “어른조사 방법론: 과제와 전망.” 대한통계협회 《통계》 33(1): 27-36.
- 허명희·한상태·강현철. 2007. 《TV시청환경 전화조사 개선방안 연구》 방송위원회 연구보고서, 지정 2007-01.
- 허명희·김영원. 2008. “RDD 표본 대 전화번호부 표본: 2007년 대통령 선거 예측 사례” 《조사연구》 9(3): 55-69.

- 허명희·황진모. 2006. “전화조사를 위한 시간균형할당표본추출.” 《조사연구》 7(2): 39-52.
홍성준·박소형·김선웅. 2009. “국내 유선 전화조사에서 표본추출틀의 포함률.” 《조사연구》
10(1): 33-56.

<접수 2011/5/31, 1차수정 2011/7/10, 2차수정 2011/7/14, 게재확정 2011/7/14>