

연구논문

임의표집법에 의거한 전화조사의 시도
-2002년 울산시장선거의 경우-

An Exploration on Random Sampling Telephone Survey
-The Case of the Ulsan Mayoral Election in 2002-

노 규 형* · 강 흥 수** · 한 철 수***

Kyu-Hyung No · Hung-Soo Khang · Cheol-Soo Han

본 논문은 임의표집법(Random Sampling)에 의거한 전화조사방법을 2002년 울산시장선거조사에 적용하여 실시한 사례를 다루었다. 임의로 표출된 1,233개 가정을 전화번호에 근거하여 생일법으로 추출한 응답자를 대상으로 최대 5회에 걸쳐 접촉을 시도하였다. 최대 5회 전화시도의 접촉 지수를 분석하여 무응답 이유 등에 대한 자진적인 평가를 실시하고, 표집된 표본집단의 인구사회학적 특성들의 일반화 비교하였으며, 조사결과와 실제 투표결과와의 비교도 다양하게 검토하였다. 또한 생일법을 통한 임의표집법의 어려움과 조사시간대 조정 및 지속적 재접촉 등 임의표집법 활성화 방안 에 대해 논의하였다.

This study reports the results of a random sampling telephone survey conducted in the case of the Ulsan mayoral election in 2002. We interview at least five times to a respondent who is randomly selected by means of the birthday method from a randomly sampled telephone number list of 1,233 households. We analyze the result of interviewing, such as absence and promise. And we compare the demographic variables of the surveyed sample and those of the population and we also compare the randomly selected sample's voting preference with outcome of the election in various ways. Finally, we discuss difficulty of random sampling with the birthday method and suggest some technical tips to conduct random sampling telephone survey.

* (주)리서치 앤 리서치 대표이사
 ** (주)리서치 앤 리서치 정치/사회조사본부장
 *** (주)리서치 앤 리서치 사회조사팀장

I. 서론

임의표집설계(random sampling design)에 기반을 두고 추출된 임의표본 이어야 표본오차의 산정과 신뢰도 측정 등 각종 통계적 처리가 가능하다는 점에서 흔히 임의표집법은 여론조사의 대표성과 신뢰성을 담보하는 가장 중요한 과학적 표본추출방법으로 여겨지고 있다. 임의 표본추출방법은 응답자를 선정할 경우에 임의성이 확실히 보장되도록 통계적 절차를 엄밀히 시행할 것을 요구한다. 만일 응답자 선정에서 임의성이 보장되지 않으면 이전 표본추출단계가 아무리 체계적으로 이루어졌다 하더라도 임의표집을 하였다고 할 수 없다. 선거조사의 경우를 예로 들자면 조사가구로 선정된 가구에 거주하는 유권자 중 어느 한 사람을 임의로 조사대상자로 선정하여 반드시 그 사람을 대상으로 조사가 진행되도록 한다. 이때 흔히 쓰이는 방법이 생일법으로 가구원 중 조사 당일에서 생일이 가장 가까운 사람을 조사대상으로 한다. 만일 선정된 가구가 조사를 거절하거나 조사대상자로 선정된 사람이 부재중일 경우에는 약속을 하고 재차 방문하여 해당되는 사람을 조사하도록 하고 만일 조사가구나 조사대상자가 거절을 할 경우에도 쉽게 다른 가구나 사람으로 대체하지 않고 몇 차례의 재시도를 하도록 한다.

실제 미국의 여론조사기관에서 실시하는 여론조사는 임의 추출된 전화 번호 목록을 이용하여 재시도, 응답 권유 등의 과정을 거쳐 조사를 실시함으로써 임의성을 엄밀히 지키는 것으로 알려지고 있다. 또한 일본의 신문사에서 실시하는 여론조사도 임의로 추출된 응답자 목록을 기초로 다수의 면접원을 단기간에 투입하여 조사를 실시하되 응답자 대체가 불가능하도록 하여 임의성을 엄밀하게 지키는 것으로 알려지고 있다.

그러나 현실적으로 우리나라 여론조사에서는 임의 추출된 응답자 목록에 의한 조사를 실시하기란 매우 어렵다. 몇 가지 이유가 있는 바,

- 유권자 목록이나 주민 목록이 외국과는 달리 공개되지 않기 때문에 표본 면접 대상자를 사전에 추출하기가 난감하다.
- 조사 기간과 비용이 대체로 제한되어 있기 때문에 현실적으로 재시

도, 설득 등 시간과 비용이 많이 드는 임의표본 추출법을 엄밀히 지키기 어렵다.

- 언론기관으로부터 의뢰 받아 조사를 수행하는 경우 짧은 시간 내에 대규모 조사를 실시하기 때문에 임의표집법을 실시하기가 매우 어렵다.

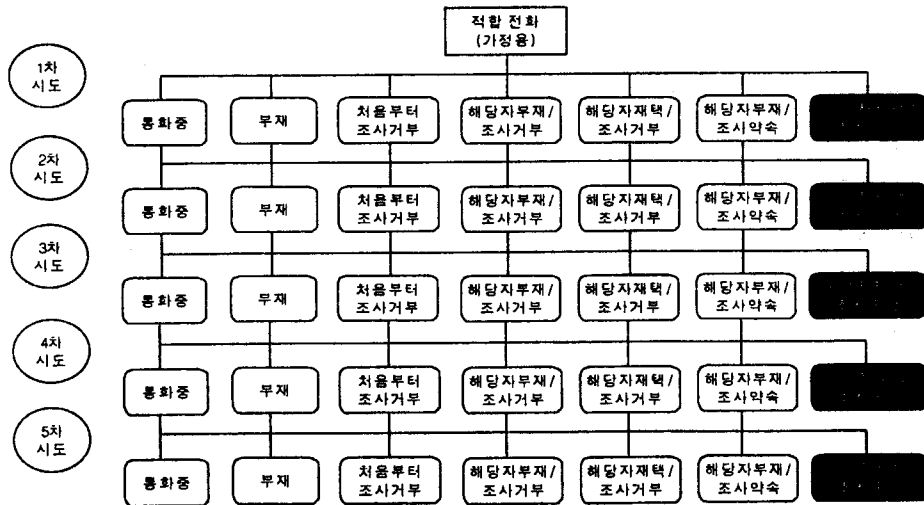
이러한 이유 때문에 우리나라 여론조사에서 가장 많이 사용되는 표본추출방법은 할당표집법이다. 즉 응답자의 지역, 성, 연령에 따른 표본크기 할당을 채우며 만일 할당을 채우지 못할 경우에는 모집단에 비례한 가중치를 주어 결과를 보정하는 방식을 택한다. 그러나 할당표집법은 임의표본추출이 아니므로 표본오차나 신뢰도 산정 등 제반 통계적 처리와 해석이 근본적으로 타당치 않음은 잘 알려져 있는 사실이다. 이 점을 인정하더라도 할당표집법의 또 다른 문제점은 응답자 대체가 쉽게 이루어지기 때문에 응답자들의 편향으로 조사결과가 틀릴 수 있다. 1948년 미국갤럽이 대통령 선거조사에서 승리자를 틀리게 예측한 것이나 1992년 영국총선에서 승리당을 잘못 예측한 경우도 표집방법이 임의표집법이 아닌 할당표집법이었기 때문이라는 지적이 있다(Moore, 1992; 허명희, 1997).

임의표집법과 할당표집법 모두 선거조사에서 많이 쓰이는 표집방법이지만 우리나라 선거조사에서는 아직까지 본격적인 임의표집법에 의한 조사가 실시된 사례가 없었음이 현실이다. ㈜리서치 앤 리서치(R&R)는 2002년 6월 13일 치러진 지방선거 직전 울산시에 대해 엄밀하게 임의추출법에 의한 조사를 실시하였다. 그 실시과정과 조사결과를 자세히 기술함으로써 임의추출법의 타당성을 검토하고자 한다. 물론 이 한 가지 사례로 임의표집법의 효과를 결론짓기는 어렵지만 그렇더라도 차후 유사 연구를 위한 기준점조사(benchmark survey)가 될 수 있을 것이다.

II. 조사 실시

선거일을 이틀 앞둔 6월 11일부터 2일간 조사를 실시하였다. 1일차(11

일) 조사는 오후 3시에서 9시까지, 2일차(12일) 조사는 오전 11시부터 오후 9시까지 진행하였다. 경험 많은 조사원 17명을 엄선하여 이들에게 임의표집법에 대해 1시간 정도 교육을 실시하였다. 조사원 교육에서는 먼저 각 조사원에게 주어진 전화번호로 전화를 걸어서 상대방과의 통화상태를 기록하되 해당 전화번호가 통화 중인지, 아무도 안 받는지를 적도록 하고, 또 여론조사기관임을 밝히고 그 집에 거주하는 유권자수를 물은 후 그 중 가장 최근에 생일이 지난 사람을 조사대상자로 정하도록 하여 반드시 그 사람을 조사하도록 하였다. 만일 해당자가 현재 집에 없으면 약속시간을 정하고 이를 응답지에 기록한 후 약속시간에 다시 전화를 걸어 조사하도록 하였다. 해당 전화번호가 통화중이거나 아무도 안 받는 경우에는 다시 전화를 걸되 최대 5차에 걸쳐 전화하도록 하고 조사에 응하지 않을 경우에는 설득을 하였다. 그래도 거부할 때는 응답사례금을 지급하겠다는 밝힘으로써 조사에 응하도록 유도하기로 하였다.



<그림 1> 조사과정 흐름도

참조: “해당지재택/조사수행”의 경우를 제외하고 나머지 사례(조사거부 포함)에서 5차까지 거듭하며 지속적으로 조사를 시도함

최초 조사인원을 통상적인 95% 신뢰도에서 오차한계 ±3 정도를 보장하는 1000명으로 하기로 하고 2002년 6월 현재 가장 최신의 전화번호부를

이용하여 전화번호 1000개를 임의 추출하였다. 그러나 첫날 조사에서 최신 인명부 전화번호를 사용하였음에도 불구하고 결번, 팩스, 직장에서 사용하는 전화번호가 다수 발견되어 전화번호를 추가로 임의 추출하여 조사를 진행하였다. 최종적으로 조사에 사용된 전화번호는 1337개였다.

<표 1>에는 추출된 전화번호 1337개에 대한 전화조회 결과를 요약하였다. 전화번호가 결번인 것이 4%, 팩스로 사용하여 전화통화가 불가능한 것이 1%, 회사나 공장 및 식당 등으로 가구가 아닌 업무용으로 사용되는 전화가 2.7%로 전체 전화번호 중 조사대상으로 부적합한 것이 7.7%를 차지하였다. 우리나라 전화는 가정용과 업무용으로 구분되어 있으나 가정용으로 신청한 전화 중 일부는 소기업이나 자영업에서 업무용으로 사용하는 경우가 있다. 참고로 지금까지 선거여론조사에서는 가정용 전화번호만을 대상으로 하고 있다.1)

필자와 R&R이 1995년에 실시한 전화조사 응답율분석에서 결번이 18.2%에 달하였던 것과 비교해본다면 금번 조사에서는 결번율이 4%에 지나지 않아 전화번호리스트의 정확성이 매우 높아졌음을 것을 알 수 있다(노규형 외, 1997). 그 이유는 1995년에는 2년마다 갱신되는 인쇄본 전화번호부를 사용하였지만 이번 조사에서는 전화번호 데이터베이스를 이용함으로써 번호정확성이 향상되었기 때문으로 사료된다.

<표 1> 사용된 전화번호의 적합성 여부

		사례수	%
적합	가정용	1,233	92.3
부적합	결번	54	4.0
	팩스	14	1.0
	업무용	36	2.7
전체		1,337	100.0

1) 조사현장에서는 소지역 조사일 경우 가정용 전화번호 이외에 업무용 전화번호를 이용하기도 하는데 이 경우 반드시 응답자가 지역내 유권자인지를 확인해야 한다.

부적합 전화번호를 제외하고 최종적으로 조사에 활용된 전화번호는 1233 사례였다. 1233개 전화번호에 대하여 최소 1회(즉 처음 전화에서 해당자를 찾고 해당자 조사를 성공한 경우)에서 최대 5회까지 접촉을 시도하여 이를 동안 임의표본추출 조사법에 의하여 전화를 건 최종회수는 모두 3751번이었다. 전화를 건 결과를 <표 2>에 제시하였다.

<표 2> 전화통화 결과

	사례수	%
전체	3,751	100.0
통화중	93	2.5
부재	1,874	50.0
처음부터 조사거부	680	18.1
해당자부재/조사거부	223	5.9
해당자재택/조사거부	68	1.8
해당자부재/조사약속	328	8.7
해당자재택/조사수행	485	12.9

총 통화 결과 2.5%는 전화가 통화중이었고(통화중), 50%는 전화받는 사람이 없었으며(부재중), 18%는 전화를 받은 후 조사자체를 처음부터 거부한 경우(처음부터 조사거부)였다. 또 5.9%(해당자부재/조사거부)는 조사대상자 즉 ‘생일이 가장 가까이 지나간 사람’이 조사 당시 집에 없지만 전화를 받은 사람이 당사자가 조사에 응하지 않을 것이라고 차단하는 경우를 말한다. 대체로 부인이나 어머니가 전화를 받으면서 남편이나 자식이 전화조사를 하지 않을 것이라고 하여 더 이상 조사를 진행할 수 없는 경우이다. 또 1.8%(해당자재택/조사거부)는 조사대상자가 집에 있지만 조사를 거부하는 경우이다. 해당자가 집에 없어서 조사약속을 하고 추후 재조사하기로 한 경우는 전체의 8.7%에 달한다(해당자부재/조사약속). 마지막으로 전체 통화수의 12.9%(해당자재택/조사수행)는 해당자가 집에 있으며 조사에 응하여 조사가 실시된 경우이다.

조사에 있어 가장 큰 걸림돌은 전화를 걸었을 때 아무도 없는 경우로서

전체 통화의 50%를 차지하고 있다. 조사가 오전 11시부터 오후 9시까지 실시되었는데도 전화시도수의 50%가 부재중이므로 전화에 의한 조사가 실제로는 많은 한계를 지니고 있음을 다시 한번 확인할 수 있다. 더구나 이들 부재자를 '집에 있는 다른 응답자로 대체'하는 할당표집법은 표본의 대표성을 훼손할 여지가 많음을 우려할 수 있다. 특히 울산지역은 공장근로자가 많고 이들 중 상당수가 젊은 층이라고 생각할 때 미혼 근로자의 경우는 조사대상자로 채택될 가능성이 매우 낮을 것이다.

III. 재접촉 결과²⁾

1. 부재중의 경우

할당표집법에 의한 전화조사에서는 응답자가 통화중이거나 부재중 또는 조사거부일 경우에 새로운 전화번호로 전화를 다시 건다. 그러나 엄밀한 임의추출법에서는 이들 전화번호에 재시도를 요구한다. 이번 조사에서는 주어진 전화번호를 4번까지 재시도하여 모두 5번까지 접촉하도록 하

<표 3> 재시도 결과

	2차시도(%)	3차시도(%)	4차시도(%)	5차시도(%)
전체	601(100)	454(100)	356(100)	255(100)
통화중	12(2.0)	5(1.1)	3(0.8)	2(0.7)
부재	372(61.8)	296(65.1)	238(66.7)	160(62.7)
처음부터 조사거부	66(11.0)	67(14.8)	50(14.0)	55(21.5)
해당자부재/조사거부	22(3.6)	12(2.6)	14(3.9)	9(3.5)
해당자재택/조사거부	6(1.0)	5(1.1)	8(2.2)	3(1.1)
해당자부재/조사약속	50(8.3)	19(4.2)	15(4.2)	5(2.0)
해당자재택/조사수행	73(12.1)	51(11.2)	28(7.8)	21(8.2)

2) 응답 기계의 오기 또는 누락 등으로 재접촉 시도에서 각 차수별로 누락한 전화번호가 소수 있었음.

였다. <표 3>에 첫 통화에서 전화를 받는 사람이 없어 부재중이었을 경우 재접촉을 시도한 결과를 정리하였다.

먼저 부재중인 601사례를 2차 시도한 결과 여전히 전화를 받지 않는 계속 부재중이 61.8%였고 처음부터 조사거부자는 11%, 해당자가 부재중이면서 전화를 받은 사람이 조사를 거부한 경우가 3.6%, 해당자가 있으면서 조사를 거부한 사람은 1%, 해당자가 부재중이어서 조사약속을 받은 사람은 8.3%였고 해당자와의 조사에 성공한 경우는 12.1%였다. 2차, 3차, 4차, 5차에 걸쳐 부재중의 비율은 평균 64%를 차지하여 가장 많았다. 해당자 부재/조사거부나 해당자재택/조사거부도 1~2%정도로 비슷하다. 다만 처음부터 조사거부가 2차, 3차, 4차에서는 11~14%정도였다가 5차에서는 20%를 넘는 것은 잦은 전화에 응답자가 거부한 것을 나타낸다. 주목할 점은 2차 시도에서 조사해당자와의 조사성공율이 12.1%, 3차시도에서는 11.2%로 10%대를 유지하고, 4차와 5차에서는 8% 대를 유지하여 재시도의 경우 조사성공율이 새 전화번호를 거는 확률과 비교하여 그리 떨어지지 않는다는 것이다. 따라서 부재중인 전화번호의 경우라도 적어도 2차례의 재시도를 할 경우 조사의 비용을 크게 늘리지 않으면서도 대표성을 높일 수 있다.

2. 약속의 경우

할당표집에서는 해당자가 집에 없을 때 전화번호를 대체하지만 임의표집에서는 약속을 통해 재시도를 함으로써 무분별한 대체를 막는다. 본 조사에서는 해당자가 집에 없을 때 전화를 받은 사람에게 해당자가 집에 있을 시간을 묻고 이들과 조사약속을 하여 재접촉을 시도하였다. <표 4>에 그 결과를 나타내었는데 성공율은 비교적 높았다. 전체 평균성공율이 12%이며 부재자의 경우도 11~8%정도인 것에 비해 평균 22%의 높은 성공율을 보이고 있다. 약속을 한 경우에는 20% 대의 높은 성공율을 보이므로 조사비용의 큰 증가 없이 표본의 대표성도 높일 수 있어 효율성이 높다.

<표 4> 조사약속을 한 경우 재접촉 성공율

	2차시도	3차시도	4차시도	5차시도
성공율(%)	20.2	25	14	29
시도수(회)	124	96	50	31

3. 시간대별 조사상황 분석

<표 5>에 주어진 응답시간대별 응답율을 보면 전화조사는 저녁이 가까울수록 부재율도 적다. 조사일이 평일인 화요일과 수요일이었는데 오전시간대인 11시경에는 62.5%가 부재였다가 오후에 들어서는 40%대로 부재율이 줄고 저녁 시간대에는 부재율이 낮아진다.

<표 5> 응답시간대별 응답율

시간대	11	12	13	14	15	16	17	18	19	20
부재율(%)	62.5	53.1	55.0	56.5	52.7	49.1	44.1	44.0	42.2	39.8
성공율(%)	10.0	11.1	10.0	10.2	10.2	9.9	11.3	18.6	17.7	17.6
시도수	129	350	360	246	569	648	408	323	491	319

부재율이 낮아짐에 따라 조사의 성공율도 높아진다. 특히 본 조사에서는 생일법에 따라 조사대상자를 임의 선정하였으므로 가족이 함께 모이는 저녁시간대에야 비로소 성공율이 높아진다.

통계청이 발표한 우리나라 생활시간 조사(통계청, 2000)에 의하면 16시경 우리나라 남자의 재택율은 22.7%, 20대의 재택율은 25%에 불과하고 17시대에는 남자는 29%, 20대는 36%에 지나지 않는다. 따라서 임의추출법을 사용할 경우 조사시간대의 선정이 매우 중요하며 그 경우 필히 휴일이나 저녁시간대를 맞추어야 할 것으로 보인다.

그러나 실제 선거조사에서는 조사결과의 적시성이나 선거변수의 영향 등을 고려할 때 이러한 조사방법상의 제약을 극복하기에 현실적으로 어려운 점이 많다.

IV. 조사 결과

최종 집계된 응답자는 485명이었고 유효 전화번호로 추출된 1233가구에 대한 응답성공율은 39.3%로 집계되었다. 전화시도회수 3751번에 대한 응답성공율은 12.9%였다.

1. 인구사회학적 특성

성별로 보면 여자가 62.7%, 남자가 37.3%로 여자가 압도적으로 많이 표집되었고, 연령별로 보면 20대가 저표집되었고 50대 이상이 과표집되었다. 소지역별로 보면 중구를 제외하고는 별도의 표집틀을 제시하지 않았음에도 불구하고 큰 차이를 보이지 않았다. 본 조사는 생일법에 의하여 조사응답자를 선정하는 절차를 거쳤음에도 불구하고 부채중이나 조사자의 거부 때문에 조사로 표집된 인구사회학적 변인이 모집단에 비해 여성과 고연령층 등 재택율이 높은 집단이 추출되어 모집단과 차이가 많다.

〈표 6〉 응답자의 인구사회학적 특성

	조사결과	모집단	차이
〈성별〉			
남성	37.3	50.1	-12.8
여성	62.7	49.9	+12.8
〈연령〉			
20대	13.2	24.8	-11.6
30대	28.9	33.0	-4.1
40대	25.4	22.6	-2.8
50대이상	32.6	19.6	+13.0
〈지역〉			
중구	18.4	23.0	-4.6
남구	32.2	31.6	+0.6
동구	20.6	17.8	+2.8
북구	10.9	11.2	-0.3
울주군	17.9	16.4	+1.5

2. 선거 결과와의 비교

이렇게 표집된 조사대상자를 근거로 조사의 목적인 선거결과를 예측한 결과를 <표 7>에 제시하였다. 편의상 당선자와 차점자만을 비교하였고 조사에서 나타난 무응답층을 제외한 응답자 기준 백분비로 나타난 결과가 실제 선거결과와 당선자의 경우 0.7% 포인트, 차점자의 경우 0.9% 포인트로 예측되어 95% 신뢰도에서 표본오차내의 매우 높은 정확성을 보였다. 앞서 지적한 대로 본 조사의 표본은 모집단과 인구사회학적 변인에서 많은 차이가 있음에도 불구하고 투표결과와 유사한 이유는 무엇일까.

투표결과를 예측하는 조사에서 조사대상이 되는 모집단은 전체 유권자가 아니라 실제로 투표에 참여하는 투표자가 되어야 마땅하다. 본 조사의 투표의향결과가 실제득표율과 일치하는 것은 아마도 본 조사의 표본이 실제 투표자를 대표하는 표본이었기 때문일 수 있다. 즉 투표율이 52.3%에 지나지 않는 이번 선거에서 여론조사에 참여한 사람들이 선거에 관심이 높고 결국 투표에 참여하는 사람들을 잘 반영한 것이라고도 할 수 있다.

<표 7> 선거예측결과

(단위: %)	선거결과	단순결과	투표율 반영 후 결과	유권자기중치 고려 후 결과	판별분석과 투표율반영 결과
박맹우 (한나라당)	53.1	53.8	55.3	45.5	49.5
송철호 (민주노동당)	43.6	42.7	41.5	50.4	45.7

☞ 판별분석시 활용변수 : 성, 연령, 학력, 직업, 소득, 출신지, 정당지지도 (판별력 : 76.9%)

유권자중에서 투표자를 사전에 알 수 없기 때문에 투표결과를 예측하는 하나의 방법으로서 각 연령집단의 과거 투표율을 반영하여 득표율을 예측하게 된다. 예상투표율을 반영한 조사결과를 비교한 결과 실제득표율과 2.1% 포인트 내외에서 적중하고 있다.

한편 모집단을 유권자 전체로 상정할 경우 득표예측은 크게 어긋난다.

특히 이번처럼 투표율이 52.3%처럼 낮을 경우에는 실제 선거결과와 예측은 크게 다르다. 본 조사결과에서 유권자 전체를 기준으로 가중치를 준 경우의 투표 예측치는 실제 선거결과와 7% 내외의 오차가 나타나고 당선자 예측마저 틀리는 오류를 보이고 있다. 또 후보지지 의사를 밝히지 않은 유보층을 판별분석하고 투표율을 반영한 결과도 실제 선거결과와 많은 차이를 보이고 있다. 즉 유보층의 지지의사를 지지를 표명한 층으로부터 유추할 수 있는 것이 아님을 보여주고 있다.

3. 선거 사후 조사

선거가 있는 이틀 후인 6월 15일(토)에 선거전 조사응답자와 조사거절자를 대상으로 재조사를 실시하였다. 그 결과 조사응답자는 56.7%인 275명이 응답에 응하였고, 조사거절자는 23명만이 조사에 응하여 3.3%에 그쳤다. 조사 응답자들 중 투표여부는 <표 8>에서 보는 바와 같이 선거이전에 밝힌 투표참여강도를 대체로 반영하고 있다. 선거전 조사에서 ‘반드시 투표하겠다’는 사람 중 92%는 실제로 투표했다고 한 반면 ‘가급적 투표하겠다’던 사람 중에서는 73%만이 투표했다. 그러나 사후 조사에 응한 응답자들이 투표한 경우가 많았을 것을 감안한다면 투표의향지수를 그대로 받아들이기는 어렵다.

<표 8> 선거전 조사에서 응답자들의 투표의향에 따른 실제투표율 비교(%)

	반드시 투표	가급적 투표	투표 안하겠다
투표율	92.4	73.2	55.6
사례수(명)	210	43	18

<표 9>는 사후 조사 응답자들이 밝힌 자신들의 지지후보인데 실제 선거결과와는 달리 당선자에 대한 지지가 매우 높은 것이 주목된다.

<표 9> 사전응답자와 사전거부자의 투표결과

	박맹우	송철호	기타/모름/무응답
사전응답자(270명)	53.9%	21%	24.3%
사전거부자(23명)	34.8%	17.4%	47.8%
실제 선거결과	53.1%	43.6%	

V. 결론

임의표집법에 의한 선거여론 전화조사를 실제로 실시해본 결과 부재중이 높고 거부율이 높아 응답자를 접촉하기가 매우 어렵다는 점이 드러났다. 부재중을 해결하는 개선방법으로 재택율이 높은 휴일에 조사하거나 초저녁 시간대에 집중적으로 조사를 실시하는 방안을 고려할 수 있다. 또 조사거부거나 부재중인 경우라도 재접촉을 시도할 경우 응답율을 제고할 수 있음을 보여준다. 4차례에 걸친 재접촉으로 응답율이 30% 이상에 도달하여 평균적인 전화조사 성공율인 14~15%의 두 배에 달하는 성과를 올렸고 약속의 경우에도 20%가 넘는 응답율을 보였다.

그러나 미국에서 조사대상자를 임의 추출하는 방법으로 흔히 쓰이는 생일법은 시행상에 문제점이 있는 듯 하다. 우선 가족의 생일을 물으면 전화통화자가 거부감을 나타내며 가족 구성원의 생일에 대해 다시 생각하게 만드는 번거로움이 있어 실제 진행에 많은 애로가 있었다.³⁾ 대안으로서 가구 내 성인 남녀 수를 고려하여 응답자를 무작위로 선정하는 Kish Grid 방법 등을 고려할 필요가 있다.

임의표집법이 갖는 실질적 장점에도 불구하고 최종 조사결과 집계된 표집이 조사응답자의 거부로 인해 모집단 유권자의 인구사회학적 분포와는 매우 다르게 나타났다. 즉 재택율이 높은 여성과 노년층에 편중되어 있었

3) 본 조사에서는 응답율을 높이기 위해 응답할 경우 인센티브를 준다고 설득하였지만 인센티브에 의한 호응도는 매우 낮았고 오히려 거부감을 표하는 경우도 있었다.

다. 그럼에도 불구하고 본 조사의 결과가 선거결과를 적중하는 성과를 보인 것은 낮은 투표율을 보인 선거라는 특수한 상황 때문인 것으로 여겨진다. 따라서 본 조사와 같은 임의추출법이 선거예측에 할당표집보다 더 우월하다고 단정짓기는 어렵다. 왜냐하면 할당표집을 사용한 다른 조사결과도 비슷한 정도의 예측 정확성을 이번 선거에서 보여주었기 때문이다.

그러나 일반적으로 어렵게 여겨지는 임의추출법에 의한 조사가 나름대로 장점을 갖고 있으며 실사과정에서 새로운 특징점을 발견했다는 점이 이번 시험 조사의 성과라 하겠다. 앞으로 임의추출법을 현실적으로 효과적으로 실행할 수 있는 방안이나 조사결과에 대한 보정 방안 등이 함께 연구되어야 할 것으로 판단한다.

참고문헌

- 노규형·김학량 공저. 1997. 《선거와 여론조사》 나남출판사.
 통계청. 2000. 《생활시간조사》 통계청.
 허명희. 1997. 《사회여론조사》 자유아카데미.
 Moore, D. 1992. *The Superpollsters* (1st ed.). N.Y.: Four Walls Eight Windows.

II. 이메일 서베이의 특징 및 응답반응 영향요인

인터넷 서베이의 유형은 크게 이메일 서베이, 웹서베이, 온라인 FGI(Focus Group Interview)의 세 가지로 분류할 수 있다(<표 1> 참조). 이 세 가지 중 이메일 서베이는 가장 신속하고 간편하며 여러 계층의 사람이 이용하는 넓은 커버리지(coverage)를 장점으로 갖고 있다. 그러나 이메일 서베이는 복잡한 설문형식을 이용하기는 어렵다는 단점이 있으며, 이러한 단점으로 인하여 복잡한 설문형식에도 이용 가능한 html형식 설문지를 이메일로 발송하는 방법이 요즘 급부상하고 있다. 또 백퍼센트 html형식 기반의 웹서베이를 이용할 수도 있다. 이 방법은 설문문항의 스킵(skip)이나 무작위화(randomization)가 가능하고 그리드 형식(grid-style) 같

<표 1> 인터넷 서베이의 유형

종 류	내 용	
이메일을 통한 서베이	Plain 이메일	조사대상자는 ASCII 파일형태로 전송된 질문에 응답을 표시하여 재발송하는 형태로 설문의 레이아웃이 제한적이고, 자료처리가 매우 힘들다.
	Disk by 이메일	설문프로그램을 이메일로 전송하면 이 파일을 다운로드하여 실행한 뒤 응답을 표시하여 파일을 재발송하는 형태로 파일의 용량과 첨부파일로 인한 바이러스의 거부감으로 거의 사용되지 않는다.
	html form-based	html형식으로 구성된 설문지를 이메일전용프로그램을 통해 이메일로 조사대상자에게 발송하는 방식으로 html의 장점을 가진다.
html에 기반한 서베이	Plain html	조사대상자에게 전체 설문문항이 한 페이지로 된 것을 제시하므로 응답을 빠트릴 수 있으므로, 짧고 단순한 설문에 적당하다.
	CAWI	Computer Assisted Web Interview는 설문의 이동, 회전, 혼합, 복수페이지로 모든 응답통제가 가능하여 가장 널리 활용되고 있다.
	Java 기반 웹설문	Java와 Active-X를 사용하여 서버와의 상호작용 없이 클라이언트에서 수행하지만 너무 느려 html형태와 선택적으로 조합하여 사용한다.
온라인을 통한 FGI	공간적 제약이 없고, 익명성을 통해 활동적 참여가 가능하나 응답의 신뢰성을 확인할 수 없다는 문제점이 있다.	

자료원: Batagelj, Lozar, and Vehovar(1998), CASRO(1998), Comley(1998), 김기수(2000) 재정리.

여 이루어지는 인터넷 서베이는 1980년대 초 특정집단을 대상으로 한 인트라넷을 이용한 전자우편 조사를 통해 그 활용 가능성이 처음으로 입증되기 시작하였다. 그러다가 1994년 이후 WWW(World Wide Web) 사용이 확대되면서 인터넷 이용자가 획기적으로 증가하고 인터넷 기술도 급격하게 발달하여 인터넷 서베이의 환경이 개선됨에 따라, 저렴하고 신속하게 응답을 얻을 수 있는 인터넷 서베이가 크게 주목받게 된 것이다.

우리나라의 경우 인터넷 서베이가 사용된 것은 언론사와 기업체에서 포털사이트(portal site)를 이용하는 방법으로 1997년 3월 KISSNET을 사용한 것이 효시이며, 현재 각 언론사와 방송사 그리고 여론조사기관에서 이미 활용하고 있거나 본격적인 준비를 하고 있다. 우리나라 인터넷 조사시장 규모가 2001년에 70억 원 정도로 추정되었고(매일경제 2000), 인터넷 사용인구가 70% 정도가 되면 인터넷을 이용한 조사가 본격적인 설문조사 방법으로 각광받게 될 것이라는 전문가의 주장(박무익 2000)도 있는 점을 감안할 때 우리나라의 인터넷 서베이 활용은 어느 정도 그 여건이 성숙되어 있다고 볼 수 있다.

인터넷 서베이는 여러 형식을 취할 수 있지만 그 중 가장 대표적인 것은 이메일 서베이이다. 이메일 서베이는 매우 신속하고 저렴하다는 장점을 갖고 있으며, 더구나 이메일 사용인구의 폭발적인 확대를 감안할 때 중요한 자료수집방법으로서의 그 잠재적 가치는 대단하다. IDC(2001)는 전 세계 이메일 계정수가 2000년에 5억500만 개에서 2005년에는 12억 개로 급증할 것이라고 전망하고 있다. 한편 우리나라의 경우에도 2000년 12월 현재 인터넷 이용자의 85%가 이메일을 보유하고 있어 인터넷 이용자의 거의 대다수가 이메일을 사용하고 있는 실정이다(Eyecllick 2001).

그런데, 이처럼 이메일 서베이 조사방법의 환경이 급속하게 변하고 있음에도 불구하고 이메일 서베이의 응답에 영향을 주는 요인에 대한 학문적 연구는 이제 출발에 불과한 형편이다(Cook, Health, and Thompson 2000). 본 연구는 먼저 이메일 서베이의 특징과 장단점 및 응답반응 영향요인을 살펴보고, 이메일 서베이 응답반응에 대한 기존의 실증연구 결과를 정리해 봄으로써 이메일 서베이에 관한 연구의 현 위치와 미래 방향을 조명해보고자 한다.