

선거여론조사의 객관성 · 신뢰성
제고를 위한 조사방법론 개선방안
연구

2019. 10.

연구기관 : 한국통계학회

본 연구 보고서는 2019년도 중앙선거여론조사심의위원회 연구
용역과제로서 연구내용은 중앙선거여론조사심의위원회 공식견
해가 아님.

연구진

박 인 호 (부경대학교 통계학과 교수)

임 종 호 (연세대학교 응용통계학과 교수)

박 민 규 (고려대학교 통계학과 교수)

차 례

제 1 장	서론	1
1.1	연구배경 및 필요성	1
1.2	총조사오차와 선거여론조사	2
1.3	연구 구성	4
제 2 장	조사방법론에 대한 통계이론 연구	6
2.1	표본추출법	6
2.2	전화조사를 위한 할당추출법	8
2.3	소결	10
제 3 장	선거여론조사의 다차원적인 비교 연구	11
3.1	서론	11
3.2	중앙선거여론조사심의위원회 등록 여론조사 기초 분석	11
3.3	심화 분석	21
3.4	소결	29
제 4 장	실험연구를 통한 표본추출틀 및 조사모드의 체계적 차이 검토	31
4.1	실험연구 개요 및 조건	31
4.2	실험연구 진행	31
4.3	실험연구 비교분석	33
4.4	실험결과 요약	48
제 5 장	선거여론조사 개선 방안 논의	50
5.1	선거여론조사 방법의 현황과 문제	50
5.2	선거여론조사 방법의 개선방법 모색	61
부록 1.	ARS 조사용 설문지	71
부록 2.	전화면접 조사용 설문지	73

표 차 례

<표 1-1> 조사통계 품질관리보고 포함요건	2
<표 3-1> 여론조사 데이터 주요 변수 요약	12
<표 3-2> 상위 10개 조사기관 여론조사 실시 현황	13
<표 3-3> 조사대상(지역) 현황	14
<표 3-4> 조사내용 현황	14
<표 3-5> 조사방법 분포	15
<표 3-6> 조사방법 수 × 조사내용	16
<표 3-7> 주된 조사방법 분포	16
<표 3-8> 단독조사 분포	17
<표 3-9> 조사내용 × 조사방법	17
<표 3-10> 혼합조사 분포	18
<표 3-11> 혼합조사 분포 × 조사내용	18
<표 3-12> 혼합조사 분포	19
<표 3-13> 조사기간 분포	19
<표 3-14> 조사방법 × 조사기간 분포	20
<표 3-15> 조사기간 - 조사방법 개수	20
<표 3-16> 조사방법별 접촉률과 응답률 (단독조사)	23
<표 3-17> 조사방법별 주요조사기관의 조사횟수 및 조사방법 비교	24
<표 3-18> ARS-A 업체 조사 결과 요약	26
<표 3-19> ARS-B 업체 조사 결과 요약	26
<표 3-20> TEL-A 업체 조사 결과 요약	27
<표 3-21> TEL-B 업체 조사 결과 요약	28
<표 4-1> 실험조건별 목표 응답수	32
<표 4-2> 실험연구를 위한 선거여론조사 업무별 일정	32
<표 4-3> 설문문항	32

<표 4-4> 최종 실험조건 구성 비교	33
<표 4-5> 실험조건별 정치관심도 비교	34
<표 4-6> 실험조건별 정치적 성향 비교	35
<표 4-7> 실험조건별 여론조사 효용성 비교	36
<표 4-8> 실험조건별 대통령 국정수행 평가 비교	37
<표 4-9> 조사평가를 위한 조사결과분류	38
<표 4-10> 실험조건별 응답분류	41
<표 4-11> 실험조건별 응답분류	42
<표 4-12> 실험조건별 산출물	43
<표 4-13> 산출물 정의	44
<표 4-14> 여심위에 등록된 최근의 접촉물 및 응답물 사례	44
<표 4-15> 전화면접조사 가상번호 응답성향모형 적합결과	46
<표 4-16> ARS조사 가상번호 응답성향모형 적합결과	47
<표 5-1> 유선번호 RDD 권역별 국번수	52
<표 5-2> 전화면접조사 가상번호 제공정도 일치율 및 추출배율	57
<표 5-3> ARS조사 가상번호 제공정보 일치율 및 추출배율	57
<표 5-4> CNN 조사연구의 품질 투명성 평가 질문	59
<표 5-5> 중앙선거여론조사심의위원회 등록 및 공표자료	60
<표 5-6> 조사통계 총조사오차 및 잠재적 영향 파악을 위한 체크리스트(예시)	67
<표 5-7> 조사통계의 지속적 품질향상을 위한 필수요건	68

그림 차례

<그림 1-1> 조사주기 관점의 총조사오차 구조	4
<그림 3-1> 유·무선전화번호 접촉률	22
<그림 3-2> 유선/무선전화번호 응답률 비교	22
<그림 3-3> 주요 여론조사 기관들의 구글 트렌드 키워드 5년간 시계열 추이 ..	24
<그림 3-4> 조사방법별 주요 여론조사 기관의 후보간 지지율 추이	25
<그림 4-1> 실험조건별 정치관심도 비교	34
<그림 4-2> 실험조건별 정치적 성향 비교	35
<그림 4-3> 실험조건별 여론조사 효용성 비교	36
<그림 4-4> 실험조건별 대통령 국정수행 평가 비교	37
<그림 5-1> 선거여론조사 조사목적에 따른 목표모집단 비교	50
<그림 5-2> 유무선 RDD 구성	53

제 1 장 서 론

1.1 연구 배경 및 필요성

선거여론조사는 민주주의의 꽃이라 불리는 선거에 대한 유권자의 전반적인 의견을 수집하는 도구로 매우 중요한 역할을 한다. 따라서 선거여론조사에 사용되는 다양한 조사방법론들의 객관성 및 신뢰성에 대한 다각적인 연구가 요구된다. 최근 들어 1인 가구 증가, 개인정보보호에 대한 인식 변화, 선거여론조사를 포함한 다양한 종류의 여론 및 마케팅 조사의 증가 등 조사환경이 날로 악화되고 있어, 이에 대응한 조사방법론 및 조사단가 적절성에 대한 연구가 필요하다. 특히 2020년 4월로 예정된 제 21대 국회의원 선거를 앞두고 공정한 선거여론조사 환경 및 필요한 조사방법론의 규범 마련을 위한 선거여론조사 방법론에 대한 연구 결과가 요구되고 있다. 통계작성방법의 전문성과 통계품질에 대한 지속적 관리 및 개선 노력은 조사통계의 정확성은 물론 질적 신뢰성에 중요한 영향을 미친다. 예로, <표 1-1>는 유럽통계위원회, 국제연합(United Nation, UN), 유럽통계청(Eurostat) 등이 1982년에 조사통계의 품질과 관련한 보고에 포함시켜야 할 최소한의 요건 제정 및 지표에 대한 지침을 요약하고 있다.

현재 「중앙선거여론조사심의위원회」에 등록된 선거여론조사들에서는 주로 할당추출(quota sampling)이라는 비확률추출법을 이용한 전화조사가 사용되고 있지만, 조사자료의 분석오차는 전통적인 확률추출(probability sampling)을 가정한 표본오차(sampling error)로만 보고되고 있다. 따라서 보다 객관적인 평가를 위해서는 조사결과에 표본오차는 물론 자료수집, 처리 및 관리 과정에서 발생하는 다양한 형태의 비표본오차(nonsampling error)도 함께 포함한 종합적인 관점에서의 조사오차를 고려하는 것이 필요하다.

표본오차는 조사대상의 일부만을 선택함에 따라 발생하는 확률적 오차(random error)로 표본이론(예, Cochran, 1977)을 적용하여 그 규모를 제어할 수 있다. 반면, 비표본오차는 조사를 기획하고 수행하는 전 과정에서 발생할 수 있다. 각 과정별로 선택되는 방법과 절차는 비표본오차의 잠재적 발생 원인이 될 수 있는데, 이는 비확

를적이고 비의도적인 오류 혹은 시스템적 과실에 해당하며 이에 대한 제어는 용이하지 않다. 최근 우리나라의 선거여론조사와 관련된 품질제고가 요구되는 상황에서, 선거여론조사에서 발생하는 표본오차와 비표본오차의 원인 및 개선방향을 모색하는 것은 매우 의미가 있다고 할 수 있다.

<표 1-1> 조사통계 품질관리보고 포함요건

(1)	자료원(data source), 정의, 분류 등과 관련된 기본정보
(2)	조사모집단 정의, 표본틀 적정성 등 조사자료 포괄범위
(3)	표본선택 및 추정방법 기술
(4)	응답률 (정의포함)
(5)	표본오차의 추정방법 및 관련측도 해석방법
(6)	발생 가능한 주요오차의 크기 및 방향은 물론 이들의 상대적 중요성 및 통계량에 미치는 영향 언급
(7)	통계량의 시계열적 비교에 영향을 미칠 조사절차의 주요변화 및 요인에 대한 언급
(8)	동일 주제에 대한 타 통계량과의 비교언급
(9)	자세하고 구체적인 기술적 절차에 대한 보고 및 참고자료 언급

1.2 총조사오차와 선거여론조사

조사결과가 모수를 얼마나 정확히 측정하는지를 나타내는 신뢰성은 총조사오차 (total survey error, TSE)의 개념으로 평가될 수 있다. 총조사오차란 표본추정량과 모수 간 제곱거리의 기댓값, 즉 평균제곱오차(mean square error, MSE)로 정의된다.

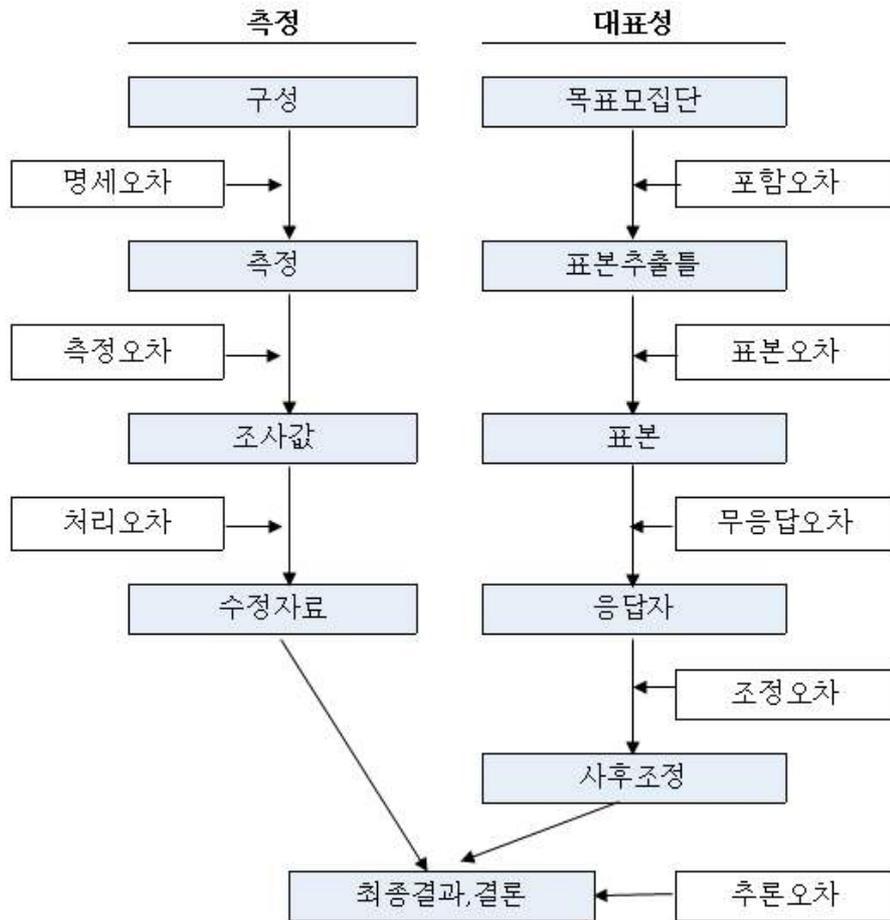
$$MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2 = Var(\hat{\theta}) + Bias^2(\hat{\theta})$$

여기서 기댓값 $E(\cdot)$ 는 반복적인 표본추출로 인해 얻게 되는 추정치의 평균값, $Var(\hat{\theta}) = E[\hat{\theta} - E(\hat{\theta})]^2$ 는 추정량 $\hat{\theta}$ 의 분산, $Bias(\hat{\theta}) = E(\hat{\theta}) - \theta$ 는 추정치의 기댓값과 모수

와의 차이인 편향(bias)을 각각 나타낸다. 매우 단순히 기술하면, 표본오차는 분산으로 비표본오차는 편향으로 나타낼 수 있다.

총조사오차는 조사설계, 자료수집, 자료처리 및 자료분석 등의 일련의 과정에서 발생할 수 있는 모든 오류를 누적한 형태의 오차(Biemer, 2010)를 뜻한다. 이는 조사품질을 결정하는 구성요소로서 조사절차의 운영적 측면에 주안점을 두는 개념이라고 할 수 있다. 총조사오차의 접근은 조사주기 내 잠재적으로 발생할 수 있는 중요한 오류요소를 식별하는 것으로 주어진 조사설계와 예산한도 내에서 조사품질을 최적화 할 수 있는 프레임틀을 제시할 수 있다. 총조사오차는 일반적으로 대표성(representation)과 관련된 오류와 측정(measurement)과 관련된 오류의 두 영역으로 구분될 수 있다. <그림 1-1>는 조사주기(life cycle of survey) 관점에서 두 영역별 세부 절차 간에 발생할 수 있는 오차를 시각화하여 나타내고 있다(Grove 외, 2009, 48쪽). 대표성 오류는 표본대상을 명시하는 표본추출틀로부터 개별 개체를 추출하는 절차, 응답과정, 조정과정 등에서 발생하는 오차들이라면, 측정관련 오류는 조사특성에 대한 개념의 설정을 바탕으로 조사문항을 구체화하고 이를 설문으로 측정하는 절차와 관련된 시스템적 오류를 뜻한다.

표본오차는 표본크기와 표본설계의 선택을 통해 모수추정의 정확성을 조절할 수 있는 반면, 그 이외의 오차인 비표본오차는 체계적 오류(systematic error)로 측정이 용이하지 않고 제거가 되지 않았을 때 특정한 방향성을 갖는 큰 편향을 가질 수 있어(Scali 외, 2005), 이에 대한 꾸준한 원인 규명과 평가가 필요하다.



<그림 1-1> 조사주기 관점의 총조사오차 구조

1.3 연구 구성

본 연구는 현재 우리나라의 선거여론조사에 현재 사용되고 있는 조사방법론에 대한 통계이론 연구, 「중앙선거여론조사심위위원회」에 등록된 선거여론조사에 대한 다차원적 비교 연구, 실험연구를 통한 표본추출틀 및 조사모드에 따른 체계적 차이를 차례로 살펴본다. 2장에서는 우리나라 선거여론조사가 기초로 하는 비확률추출인 할

당추출에 대해 통계이론에 기반한 논의를 포함한다. 3장에서는 「중앙선거여론조사 심의위원회」에 등록된 선거여론조사에 대해 RDD 기반의 유선, 무선, 그리고 유무선 병행, 그리고 가상번호 등의 서로 다른 표본추출틀의 사용에 따른 조사결과 비교, 전화조사와 ARS조사 등 조사모드에 따른 조사결과에 대한 비교 연구, 또한 응답률 및 접촉률에 따른 조사결과의 비교연구를 다룬다. 4장은 실험연구를 통한 유선 및 무선 RDD, 가상번호 등 표본추출틀에 따른 체계적 차이, 전화조사와 ARS조사 등 조사모드에 따른 체계적 차이에 대해 다룬다. 마지막으로 5장에서는 현행 선거여론조사에 적용되는 조사방법론에 대한 총조사오차 관점에서의 문제점 및 개선방향에 대해 간단히 논의하며, 더불어 가상번호 기반의 확률추출 적용의 가능성에 대해서도 검토한다.

제 2 장 조사방법론에 대한 통계이론 연구

2.1 표본추출법

유한 모집단(finite population)의 특성을 나타내는 모수(parameter) 추정이 주요 목적인 전통적인 서베이 통계학은 확률적 표본추출법(probability sampling) 혹은 확률표본설계(probability sampling design)에 그 통계 이론적 근거를 두고 있다. 즉, 전통적인 서베이 통계학에서는 통계적 추론을 위하여 필수적으로 요구되는 확률 구조(probability structure)가 확률표본설계 및 이에 근거한 표본추출을 통해 정의되고, 이를 이용하여 추정량들의 분포가 결정된다. 이는 모집단에 대한 분포 혹은 모형 가정에 근거한 추론이 이루어지는 타 통계학 분야의 접근 방법과는 근본적인 차이가 있음을 의미한다. 확률표본설계의 정의는 다음과 같다(박민규·강현철: 2016).

확률표본설계: 크기가 N 인 유한 모집단 $U = (1, 2, \dots, N)$ 의 모든 가능한 부분집합으로 구성된 집합을 \mathcal{E} 로 정의할 때 확률표본설계 $p(s)$ 는 다음의 조건을 만족하는 집합함수 $p(s): \mathcal{E} \rightarrow [0, 1]$ 로 정의된다.

$$(i) p(s) \geq 0, \quad (ii) \sum_{s \in \mathcal{E}} p(s) = 1, \quad (iii) \pi_k = \Pr(k \in s) > 0, \quad k \in U$$

통계청의 승인을 받은 조사들을 위한 표본설계로는 항상 확률표본설계가 사용되며 이는 이를 통해 작성된 통계가 모집단의 모수를 최소한 편향 없이 추정해야 하기 때문이다. 대표적인 확률표본설계로는 단순임의추출법, 층화추출법, 집락추출법, 이상추출법 그리고 언급한 여러 추출법이 함께 사용된 층화집락추출법과 같은 복합 표본설계이다. 이러한 확률표본추출법에 관한 이론적 내용은 박민규·강현철(2016)을 참고하면 된다.

비확률 표본설계 혹은 표본추출 방안은 위의 정의에서 조건 (iii)을 만족하지 않는 표본설계를 의미하며, $\pi_k = 0$ 인 표본설계하에서는 모수에 대한 비편향 추정량을 구할 수 없다. 대표적인 비확률 표본설계 방안으로는 판단추출법(judgment

sampling), 자원자 추출법(volunteer sampling) 그리고 할당 추출법(quota sampling) 등이 있다. 판단추출법과 자원자 추출법은 각각 개체 포함확률 π_k 가 조사자와 피조사자에 의하여 주관적으로 결정되며, 따라서 π_k 의 값이 0인 개체가 발생하게 되며 또한 조사대상의 선택 역시 확률적 메커니즘에 의하여 이루어지지 않아 산출된 추정량의 통계적 성질을 유도할 수 없다.

비확률 표본추출법으로 조사 현장에서 많이 사용되는 할당추출법은 여전히 그 이론적 타당성에 대한 논쟁을 일으키고 있다. Lohr(1999)는 그의 책에서 할당추출법을 아래와 같이 두 단계에 의하여 표본이 추출되는 방안으로 정의하였다.

Step 1: 모집단을 상호 배타적인(mutually exclusive) 부모집단(혹은 할당 층)으로 구분한다.

Step 2: 각 부모집단으로부터 할당된 수의 개체를 비확률 표본추출법을 통해 추출한다.

즉, Lohr의 정의에 의하면 할당표본추출법은 각 부모집단으로부터의 표본추출이 비확률적으로 이루어지는 표본설계를 의미한다. 위 정의에 의하면 선거여론조사를 포함한 대부분의 여론조사를 위하여 실제 사용되는 전화 조사의 경우에는 전화번호의 추출 절차 그리고 결번 및 부적격 이외의 무응답 처리 방안을 자세히 살펴보고 이를 기반으로 확률 표본설계 여부를 결정해야 할 것이다.

할당추출법과 유사한 추출방안으로는 균형추출법(balanced sampling) 혹은 이를 구현하는 방법을 의미하는 제한적 무작위 추출법(restricted random sampling)이 있다. 균형 표본은 모형기반 추론(model-based inference)을 위하여 가정된 모형 하에서 유한 모집단의 특성을 나타내는 모수 추정량의 효율성을 높이기 위하여 Valliant, Dorfman, Royall(2000) 등에 의하여 소개되었으며 그 정의는 다음과 같다. 주어진 함수 f 와 $\delta \geq 0$ 에 대하여

$$|f(\mathbf{X}_s) - f(\mathbf{X}_U)| \leq \delta$$

을 만족하는 표본을 균형 표본으로 정의한다. 여기서 \mathbf{X}_s 와 \mathbf{X}_U 는 표본과 모집단에
서 균형 조건을 정의하기 위해 사용되는 보조변수(auxiliary variables)의 행렬을 나
타낸다. $\delta=0$ 인 경우가 완전한 균형이 이루어진 표본을 의미하여 실제 이러한 표
본을 추출하고 조사하는 것은 매우 어렵다. 일반적으로 많이 사용되는 함수 f 는 평
균으로 위의 균형 조건은 $|\bar{\mathbf{x}}_s - \bar{\mathbf{x}}_U| \leq 0$ 로 표현된다. 만약, 보조변수가 부모집단을
구별하는 지시함수로 구성되었고 $\delta=0$ 인 경우의 균형 표본은 모집단의 부모집단
분포와 표본의 분포가 일치하는 표본을 의미한다.

제한적 무작위 추출법이 할당추출법과 다른 점은 실제 표본을 추출하는 과정에서
나타난다. Valliant, Dorfman, Royall(2000)에 의하여 제안된 제한적 무작위 추출법
의 과정은 다음과 같다.

Step 1: 단순임의추출법을 통해 크기가 n 인 표본을 추출한다.

Step 2: 균형 조건 $|f(\mathbf{X}_s) - f(\mathbf{X}_U)| \leq \delta$ 을 만족하면 추출된 표본을 최종 표본
으로 정의하고, 그렇지 않은 경우에는 step 1의 과정을 반복한다.

즉, 제한적 임의추출법에서는 모든 k 에 대하여 $\pi_k > 0$ 이며, 따라서, 확률표본추출법
의 기본 조건에 부합한다. 그러나 정확한 π_k 의 값의 계산은 가능하지 않다.
Park(2006)은 제한적 임의추출법 하에서 각 개체가 표본에 포함될 확률을 근사적
으로 유도하고 이를 이용하여 계산된 추정량이 단순임의추출법의 가정하에서 정의
되는 회귀 추정량과 근사적으로 일치함(asymptotically equivalent)을 증명하였다.

2.2 전화조사를 위한 할당추출법

여론조사를 위해 흔히 사용되는 전화 조사에서는 2.1절에서 언급된 할당추출법
이 주로 사용된다. 할당추출을 위해서는 성, 지역 그리고 나이 변수가 주로 사용된
다. 즉, 전국 단위의 지역×성×나이에 대한 모집단 분포를 이용하여 표본 각 층의
할당 수를 결정하고 배분된 수만큼의 표본을 비확률적인 방법을 이용하여 채우게

된다. 일반적으로 전화 조사에서는 전화 걸기와 통화가 이루어진 개인을 대상으로 조사가 동시에 이루어진다. 즉, 이 과정에서 전화 걸기를 일종의 표본추출로 이해할 수 있으며, 통화 후 설문을 묻는 과정을 조사로 이해할 수 있다.

조사를 위하여 사용되는 전화의 종류에 따라 전화 조사는 유선, 무선 그리고 유무선 병행조사로 구분할 수 있다. 유선전화 조사의 추출단위는 유선전화의 특성으로 볼 때 가구로 볼 수 있으며, 무선전화의 추출단위는 개인으로 볼 수 있다. 유무선 병행조사의 경우에는 서로 다른 추출단위인 가구와 개인이 조사를 위해 정의된다. 무선전화의 보급률이 낮은 과거에는 유선전화를 이용한 전화 조사의 모집단 커버리지에 문제가 없었으나 국민 대부분이 무선전화를 보유하고 있으며 유선전화를 보유한 가구의 비중이 급격히 낮아짐에 따라 유선전화만을 이용한 조사에 대한 문제가 제기되었고 이를 보완하기 위하여 무선전화 조사 혹은 유무선 병행 전화 조사가 도입되었다. 또한, 유선과 무선의 전화번호 리스트가 존재하지 않는 상황에서 가능한 모든 번호가 추출될 수 있도록 3가지 형태의 조사를 위하여 RDD(random digit dialing)가 사용되고 있다. 비용 대비 효과적인 RDD의 사용을 위해서는 기본적으로 존재하지 않는 번호(not working number) 및 부적격 번호(not eligible number)를 구분해야 하고, 무응답을 최소화하기 위한 대책이 마련되어야 한다.

만약, RDD를 통한 전화번호의 추출이 단순임의추출 표본을 얻기 위한 순차적 과정(sequential process)이며, 적절한 수의 재접촉이 이루어지고, 추후 통계적으로 타당한 무응답 조정이 이루어지면 할당추출법은 2.1절에서 언급된 균형추출법으로 이해할 수 있다. 그러나, 실제 전화 조사를 위해 조사 현장에서 사용되는 RDD가 단순임의추출표본을 얻기 위한 무작위 과정인지에 대해서는 실제 사용되는 방법에 대한 심층적인 검토가 필요하다. 특별히 현재 대부분 전화 조사에서 사용되는 RDD 기법은 유선과 무선에서는 일종의 패널이 운영되는 것이 아닌가 하는 의심이 있으며, 또한, 유무선 병행조사의 경우에는 통계 이론적으로 타당한 다중 추출틀을 활용한 표본추출이 아닌 할당 기준을 맞추기 위한 비확률추출법이 사용되고 있는 것으로 판단된다. 외국 사례를 살펴보면 유무선 병행조사를 위한 표본추출은 일종의 다중 추출틀을 이용한 방안으로 이해할 수 있다. 유무선 병행조사가 두 개의 추출틀이 이용되는 dual frame survey로 정의되기 위해서는 유선 번호와 무선 번호의 추

출을 위한 확률 표본추출법이 별도로 정의되고 이에 따른 번호 걸기와 조사가 이루어져야 하나 언급한 바와 같이 국내에서는 계획된 할당 기준을 채우는 방안으로 유무선 병행조사가 사용되고 있어 전통적인 다중 추출틀 표본설계로 볼 수 없다.

2.3 소결

선거여론조사를 위해 국내에서 사용되는 국내의 전화 조사는 조사방식에 따라 면접 조사와 ARS 조사, 조사 도구에 따라 유선, 무선 그리고 유무선 병행조사로 구분되고 있다. 그러나 언급한 조사방식 모두를 위한 표본추출 방안으로는 할당추출법이 사용되고 있으며 현재까지 파악한 바로는 이는 비확률추출법이며, 따라서 확률 표본추출에 근거를 준 전통적인 통계 이론적 측면에서 그 결과의 타당성은 확보되지 않는다.

이러한 상황에서 표본설계 측면에서 볼 때 선거여론조사 자료의 분석을 위해서 취할 수 있는 전략은 첫째로는 검증이 이루어진 RDD의 적용과 이에 근거한 디자인 기반 추론이며, 둘째로는 균형추출법의 적용 후 모형기반의 추론이다. 실제 이러한 두 가지 가능한 방안에 대한 국내의 이론적 혹은 경험적 연구는 거의 전혀 없는 상황이다. 이는 조사기관이 사용하고 있는 RDD 기법이 표준화되어 있지 않고 또한 조사결과를 사용하거나 조사를 의뢰하는 기관 역시 이 부분에 대한 전문성 부족으로 이를 요구하고 있지 않기 때문이며 이러한 부분의 개선이 없이는 통계 이론적으로 타당한 결과를 얻기 위한 최소한의 방법론적 기본을 만족하지 않은 조사가 계속해서 사용될 것이다.

더불어 최근 RDD의 대안으로 가상번호의 사용이 제안되고 있으나 먼저 해결해야 할 문제는 각 통신사가 제공하는 무선번호의 추출에 대한 무작위성 혹은 확률추출법의 적용이다. 즉, 현재와 같이 각 통신사가 제공하는 번호의 추출 방식에 대한 정보가 전혀 없는 상황에서는 이를 기반으로 한 결과의 타당성 역시 통계 이론적 측면에서 확보할 수 없을 것이다.

제 3 장 선거여론조사의 다차원적 비교연구

3.1 서론

일반적으로 여론조사는 조사대상과 내용, 데이터 수집 및 분석의 모든 과정에서 오류가 발생할 수 있다. 이러한 측면에서 최근에는 표본오차(sampling bias)와 비표본오차(non-sampling bias)를 모두 포괄하는 총조사오차(total survey error, Groves and Lyberg, 2010)의 관점에서 조사결과를 분석하고 해석하려는 사례가 늘고 있다. 본 연구에서는 이러한 개념을 활용하여 현행 선거여론조사를 다차원적으로 연구하는 것을 목표로 하였다. 하지만 조사 결과에 영향을 줄 수 있는 모든 요인들을 관측하거나 혹은 측정할 수 없기 때문에, 여론조사에 영향을 줄 수 있는 주요한 요인(factor)들을 조사기관, 조사방법, 응답률 등을 중심으로 여론조사 결과들을 비교 분석하였다.

본 장에서는 두 가지 형태의 데이터 분석을 포함한다. 3.2 절에서는 2014년 3월부터 2019년 7월 사이에 중앙선거여론조사심의위원회 홈페이지에 등록된 선거 관련 여론조사들에 대한 비교 분석을 실시하였다. 3.3절에서는 19대 대통령선거 한 달 전에 실시된 선거여론조사 결과를 이용하여 보다 심층적인 비교 분석을 실시하였다.

3.2 중앙선거여론조사심의위원회 등록 여론조사 기초 분석

1) 데이터

본 절의 데이터 분석은 2014년 3월부터 2019년 7월 중앙선거여론조사심의위원회 홈페이지에 등록된 총 5,616건의 선거 관련 여론조사를 대상으로 하였다. 분석에 사용한 주요 변수에 대한 정보는 <표 3-1>에 요약되어 있다.

<표 3-1> 여론조사 데이터 주요 변수 요약

변수	내용
조사기관	조사 실시 기관
의뢰기관	조사 의뢰 기관
조사대상	전국 단위 /지역 세부 단위
조사내용	대선/총선/지선/재보궐선거/기타
조사기간	당일/2일/3일/4일/5일/6일 이상
조사방법	유선ARS, 무선ARS, 유선전화면접, 무선전화면접 스마트폰앱조사, 기타(전화면접)
표본추출틀	유선전화번호/무선전화번호/기타
표본추출방법	RDD/무작위추출/통신사제공/기타
추출틀규모	연속형 (조사방법별)
피조사자수	연속형 (조사방법별)
비적격사례수	연속형 (조사방법별)
무응답사례수	연속형 (조사방법별)
응답사례수	연속형 (조사방법별)
가중치	가중치 없음/셀가중/립가중
표본오차	연속형

2) 변수별 데이터 분석

본 절에서는 주요 변수 중에서 조사기관, 조사기간, 표본추출틀, 조사방법, 응답률, 가중치 처리 방법, 표본오차를 중심으로 여론조사를 요약한 뒤 결과에 대한 비교 분석을 실시하였다.

가. 조사기관

총 5,616개의 여론조사는 159개의 조사기관에 의하여 실시되었다. 이 중 상위 44개 기관이 기간 내 전체 여론조사의 90%인 총 5,043회를 담당하였다. 159개의 조사기관 중에서 기간 내 여론조사 횟수가 10번 미만인 기관은 99개 업체로 이 중에서 35개 (전체 기관 수 대비 약 22%) 업체는 한 번의 여론조사만 실시하였다. 상위 10

개 조사기관은 <표 3-2>에 정리되어 있다. 상위 10개 기관의 조사 횟수는 총 3,324 회로 기간 내 전체 조사 대비 약 59.2%에 해당한다. 분석의 편의상 향후 자료 분석의 일부는 상위 10개 조사기관만을 대상으로 하였다.

<표 3-2> 상위 10개 조사기관 여론조사 실시 현황

기관명	여론조사 횟수	비율 (%)
리얼미터	1,449	25.8
한국갤럽조사연구소	292	5.2
한국리서치	281	5.0
코리아리서치센터	236	4.2
조원씨앤아이	229	4.1
리서치뷰	190	3.4
모노커뮤니케이션즈(모노리서치)	176	3.1
리서치앤리서치	170	3.0
알앤씨치	152	2.7
마크로밀엠브레인	149	2.7
합계	3,324	59.2

상위 10개의 조사업체 중에서 통계청 산하 비영리 단체인 한국조사협회(KORA)에 소속된 회원사는 5곳 (한국갤럽조사연구소, 한국리서치, 코리아리서치센터, 리서치앤리서치, 마크로밀엠브레인)이다. 한국조사협회는 조사윤리강령과 ARS 행동규범으로 인하여 2014년 6월 18일 이후로는 ARS 조사를 실시하고 있지 않는 특징이 있다. 이를 통하여, 현행 여론조사 시장이 ARS 조사에 대한 태도와 입장에 따라서 구분되어 있음을 확인할 수 있다. ARS 조사와 전화 조사에 대한 비교와 분석은 4장의 실험조사 데이터에 대한 분석을 참고할 수 있다.

상위 10개 조사업체 중에서 리얼미터가 1,449회 (25.8%)를 실시하여 가장 큰 비율을 차지하였으며 한국갤럽조사연구소가 292회로 5.2%의 비중을 차지하였다. 여기에서 특이한 점은 두 조사기관 모두 주간 동향 조사를 실시하고 있음에도 불구하고, 한국갤럽조사연구소는 177회의 자체조사가 기록되어 있으며 리얼미터는 4회만 자체조사로 등록되었다는 것이다.

나. 조사대상 및 내용

조사내용에 따라서 전국 혹은 지역 단위로 조사가 실시되었으며 전국 단위로는 32.8%, 시도 단위로는 67.2%의 조사가 실시되었다. 조사대상(지역)과 조사내용은 <표 3-3>과 <표 3-4>에 각각 요약되어 있다.

<표 3-3> 조사대상(지역) 현황

대상지역	조사횟수	비율 (%)	대상지역	조사횟수	비율 (%)
전국	1,842	32.8	서울특별시	356	6.3
강원도	143	2.5	세종시	61	1.1
경기도	625	11.1	울산광역시	67	1.2
경상남도	207	3.7	인천광역시	143	2.5
경상북도	368	6.6	전라남도	322	5.7
광주광역시	138	2.5	전라북도	294	5.2
대구광역시	184	3.3	제주도	153	2.7
대전광역시	112	2.0	충청남도	284	5.1
부산광역시	138	2.5	충청북도	179	3.2

<표 3-4> 조사내용 현황

조사내용	조사횟수	비율 (%)
제19대 대통령선거	830	14.8
제20대 대통령선거	61	1.1
제20대 국회의원선거	1,657	29.5
제21대 국회의원선거	76	1.4
제6회 전국동시지방선거	815	14.5
제7회 전국동시지방선거	1,692	30.1
2014년 상반기 재보궐선거	64	1.1
2015년 상반기 재보궐선거	45	0.8
2015년 하반기 재보궐선거	2	0.0
2016년 상반기 재보궐선거	4	0.1
2017년 재보궐선거	12	0.2
2018년 재보궐선거	27	0.5
2019년 재보궐선거	17	0.3
기타	314	5.6

<표 3-4>에서 21대 국회의원 선거 (2020년)와 20대 대통령 선거(2022년)는 아직 치러지지 않았기 때문에 조사 횟수가 많지 않으며, 기타 항목은 정당후보지지도, 국회의원 의정활동평가, 시도 평가, 차기 대선주자 여론조사 등을 포함하고 있다.

다. 조사방법

여론조사 방법으로는 ARS조사, 전화면접조사, 스마트폰앱조사, 인터넷조사, 대인 면접조사 등이 실시되고 있다. 본 보고서에서는 유무선ARS 조사와 유무선전화면접조사의 경우, 동일한 조사방법에 추출틀이 2가지인 형태이지만, 분석의 편의상 혼합조사로 간주하였다.

<표 3-5> 조사방법 분포

조사방법 갯수	조사횟수	비율 (%)
1 가지	2,129	37.9
2 가지	2,667	47.5
3 가지	410	7.3
4 가지	357	6.4
5 가지	53	0.9

※조사방법 기재여부를 조사비중보다 우선으로 평가¹⁾

<표 3-5>는 실제 조사에 사용된 조사방법의 숫자에 따른 분포를 표기하고 있으며, 1 가지의 조사방법만 사용된 경우는 약 38%이고 62%는 두 개 이상의 조사 방법이 혼합되어 사용되었다. 모집단 프레임에 따른 차이가 있을 수 있기에, ARS와 전화면접 조사의 경우, 유선 혹은 무선 한 가지로만 조사가 진행되었으면 조사 방법이 한 가지 인 것으로 분류하였으며, 유무선을 혼합해서 한 경우에는 조사 방법이 두 가지 사용된 것으로 분류하였다. 세 가지 이상이 혼합되어 진행된 경우도 약 14.6%나 되었다. 해당 여론조사를 위한 조사방법 선택에 대한 이유나 사유가 결과 등록에 기재되지 않기 때문에 심층적인 분석은 불가능하다. 대신에 본 연구에서는 조사대상(내용)과의 교차분석을 실시하였으며 이를 통하여 단독/혼합조사 선택에 따른 차이를 간접

1) 조사비중이 기재가 안된 경우 비중을 0으로 표기하였다. 조사비중합이 99-101 사이가 아닌 조사는 9개이며, 이 중에서 1개는 조사 비중 합이 0이다(홈페이지 등록번호: 227).

적으로 논의할 수 있을 것으로 기대한다.

<표 3-6> 조사방법 수× 조사내용 (세로합 100%)

조사방법 수	19대 대통령	20대 국회	6회 지선	7회 지선
1가지	118 (14.2)	863 (52.1)	576 (70.7)	396 (23.4)
2가지	321 (38.7)	654 (39.5)	229 (28.1)	1153 (68.1)
3가지	82 (9.9)	44 (2.7)	10 (1.2)	141 (8.3)
4가지	256 (30.8)	96 (5.8)	0 (0.0)	2 (0.1)
5가지	53 (6.4)	0 (0.0)	0 (0.0)	0 (0.0)

<표 3-6>을 보면 조사내용과 조사에 사용된 방법 가짓수에 대한 교차표가 표시되어 있다. 대체적으로 지역단위/전국단위 조사에 따른 차이보다 시간에 따른 혼합조사 비중의 차이가 더 두드러지는 것을 확인할 수 있다. 즉, 제19대 대통령선거와 제7회 지방선거에서 혼합조사 비중이 높은 것을 확인할 수 있다. 또한, 제6회 지방선거와 비교하여 제7회 지방선거에서는 단독조사보다 혼합조사에 대한 비중이 큰 폭으로 증대된 것을 확인할 수 있다.

전체조사에서의 주된 조사방법은 <표 3-7>에 정리되어 있다. 이 때, 조사비중이 가장 높은 조사방법을 주된 조사방법으로 가정하였다. 유선ARS가 주된 조사방법으로 사용된 조사는 약 36.1%로 전체 조사의 1/3정도가 되며, 무선ARS와 유무선전화면접 또한 높은 비중으로 사용되었음을 확인할 수 있다.

<표 3-7> 주된 조사방법 분포

주된조사방법	조사횟수	비율 (%)
유선ARS	2,027	36.1
무선ARS	1,238	22.0
유선전화면접	1,107	19.7
무선전화면접	963	17.1
기타	281	5.0

<표 3-8>은 단독조사에 사용된 조사방법의 분포를 표시해주고 있다. 여기에서, 유선ARS와 유선전화면접 비중이 85.5% 가장 높는데 이는 전체 조사를 가장 많이 실시한 조사기관의 주요한 조사방법에 기인하는 측면과 무선ARS 및 무선전화면접으로 는 데이터를 수집하기 힘든 지역조사에서는 유선ARS와 유선전화면접이 선호되는 점이 반영된 것으로 사료된다.

<표 3-8> 단독조사 분포

조사방법	조사횟수	비율 (%)
유선ARS	1,159	54.4
무선ARS	259	12.3
유선전화면접	662	31.1
무선전화면접	24	1.1
기타	25	1.2

<표 3-9> 조사내용 × 조사방법 (가로합 100%)

조사내용	유선ARS	무선ARS	유선전화면접	무선전화면접	기타
19대 대선	19 (16.1)	83(70.3)	6 (5.1)	2 (1.7)	8(6.8)
20대 국회	386 (44.7)	32(3.7)	440 (51.0)	0 (0.0)	5(0.6)
6회 지선	415 (72.0)	5(0.9)	154 (26.7)	1 (0.2)	1(0.2)
7회 지선	274 (69.2)	83(21.0)	11 (2.8)	19 (4.8)	9(2.3)

두 가지 방법을 혼합하여 실시된 조사의 경우에는 유·무선전화면접 (47.0%)과 유·무선ARS (35.5%) 조합이 가장 많았으며 해당 결과는 <표 3-10>에 요약되어 있다. 두 주요 혼합조사²⁾ 방법은 조사대상과 크게 상관없이 유효하게 사용되고 있는데 관련된 내용은 <표 3-11>을 통하여 확인할 수 있다. 다만 제6회 지방선거에서는 유·무선전화면접이 많았던 반면에 다른 세 선거에서는 유선ARS 중심의 여론조사가 많았음을 확인할 수 있다. 즉 유선ARS가 혼합조사에서 다른 조사와 함께 많이 쓰이고 있는데 여론조사에 대한 시장점유율과 밀접한 관계가 있을 것으로 사료된다.

2) 유무선ARS와 유무선전화면접의 경우도 분석의 편의상 혼합조사로 간주하였다.

<표 3-10> 혼합조사 분포

조사방법	조사횟수	비율 (%)
유선ARS+무선ARS	947	35.5
유선ARS+유선전화면접	25	0.9
유선ARS+무선전화면접	23	0.8
유선ARS+기타	305	11.4
무선ARS+유선전화면접	7	0.3
무선ARS+무선전화면접	1	0.0
유선전화면접+무선전화면접	1,253	47.0
유선전화면접+기타	104	3.9
무선전화면접+기타	1	0.0
기타+기타	1	0.0

<표 3-11> 혼합조사 × 조사내용 분포 (세로합 100%)

조사방법	19대 대선	20대 국회	6회 지선	7회 지선
유선ARS+무선ARS	98 (30.5)	85 (13.0)	39 (17.0)	596(51.7)
유선ARS+유선전화면접	1 (0.3)	15 (2.3)	3 (1.3)	2 (0.2)
유선ARS+무선전화면접	0 (0.0)	1 (0.2)	1 (0.4)	17 (1.5)
유선ARS+기타	8 (2.5)	289(44.2)	0 (0.0)	4 (0.3)
무선ARS+유선전화면접	0 (0.0)	0 (0.0)	3 (1.3)	4 (0.3)
무선ARS+무선전화면접	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
유선전화면접+무선전화면접	189(58.9)	206(31.5)	164(71.6)	530(46.0)
유선전화면접+기타	24 (7.5)	58 (8.9)	19 (8.3)	0 (0.0)
무선전화면접+기타	1 (0.3)	0 (0.0)	0 (0.0)	0 (0.0)
기타+기타	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)

혼합조사에 사용되는 방법의 수가 올라갈수록 유선ARS가 들어가는 비중이 훨씬 더 증가하는 경향이 있는데, 이는 <표 3-12>에서 확인할 수 있다. 3가지 방법이 혼합된 조사에서 유선ARS가 사용된 비중은 90.5% (=77.6%+9.0%+3.9%) 유독 높은 것을 확인할 수 있는데 비용 측면에서 효과적이고 다른 조사방법과의 혼합이 쉽기 때문인 것으로 이해된다.

<표 3-12> 혼합조사 분포

조사방법	조사횟수	비율 (%)
유선ARS+무선ARS+무선전화면접	318	77.6
유선ARS+무선ARS+기타	37	9.0
유선ARS+유·무선전화면접	16	3.9
유선+무선+무선 전화면접	16	3.9
유·무선전화면접+기타	11	2.7
기타 조합	11	2.7

라. 조사기간

조사기간은 대체적으로 3일 이내(89.5%)이며 2일 이내(65.9%)의 비중도 높은 것을 알 수 있다.

<표 3-13> 조사기간 분포

조사기간	조사횟수	비율 (%)
1일(당일)	1,048	18.7
2일	2,650	47.2
3일	1,327	23.6
4일	307	5.5
5일 이상	284	5.0

<표 3-14>를 보면 유선ARS 조사방법이 다른 조사방법에 비하여 상대적으로 당일 이내 조사 비중이 높은 것을 확인할 수 있으며, 무선보다는 유선을 기반으로 한 경우에서 2일 이내 조사의 비중이 높은 것을 확인할 수 있다.

<표 3-14> 조사방법×조사기간 분포 (가로합 100%)

조사방법	1일	2일	3일	4일	5일+
유선ARS	29.8	45.4	18.8	3.8	2.1
무선ARS	14.4	45.2	24.8	3.9	11.7
유선전화면접	13.4	52.1	26.1	7.1	1.3
무선전화면접	9.0	47.7	32.6	5.7	5.0
기타	10.7	47.3	12.5	17.1	12.5

<표 3-15>는 2일 이내로 조사가 진행된 3,799의 조사의 조사방법 개수 분포를 나타내고 있다. 비율1이 해당 분포이며, 비율2는 전체 조사에서의 분포로 <표 3-5>에서 얻어진 결과이다. 조사방법의 수가 많아질수록 전체적인 조사기간이 길어지는 것을 확인할 수 있다.

<표 3-15> 조사기간-조사방법 개수 (2일 이내 조사 한정)

조사방법 개수	조사횟수	비율1(%)	비율2(%)
1 가지	1,633	44.2	37.9
2 가지	1,715	46.4	47.5
3 가지	115	3.1	7.3
4 가지	210	5.7	6.4
5 가지	25	0.7	0.9

마. 접촉률 및 응답률

대부분의 여론조사가 성별, 지역, 연령대를 활용하여 할당추출을 하고 있다. 이러한 표본설계 구조하에서는 접촉률 및 응답률이 매우 중요할 수 있다. 특정 집단 (성별×연령×지역)에 대한 접촉과 응답률이 매우 낮은 경우, 조사 대상자에 대한 선택 편이가 발생할 수 있으며 대표성 문제가 발생할 수 있다. 또한, 시간과 비용이 증가될 수 있기 때문에 조사방법의 선택에 있어서도 선택편의가 발생할 수 있다. 이러한 측면을 논의하기 위하여 접촉률과 응답률을 정의하고 조사방법과의 관계를 분석하였다. 할당

에 따른 효과는 현재 등록 데이터로는 확인이 힘들기 때문에 추후 실험 데이터에서 추가적으로 분석이 되었다.

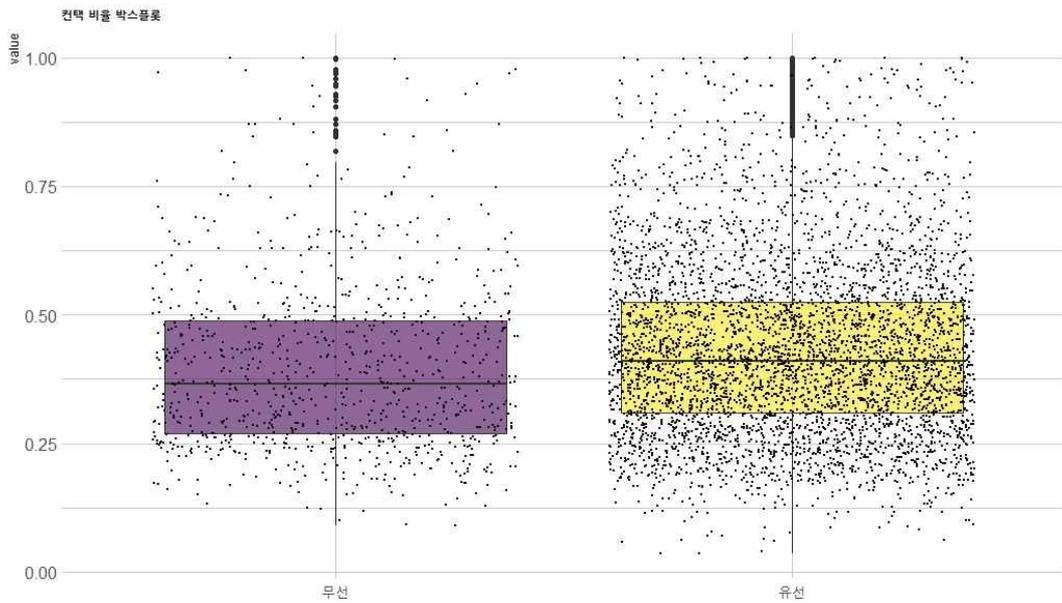
본 연구에서는 등록 자료의 성격상 유·무선 전화번호에 대한 접촉률과 응답률을 중심으로 비교하였으며, 혼합조사의 경우 주된 방법으로 판단된 방법을 중심으로 추정을 구하였다. 접촉률과 응답률은 다음과 같은 정의를 이용하여 계산하였다.

접촉률: (접촉 성공 사례 수)/(총피조사수-비적격사례수)

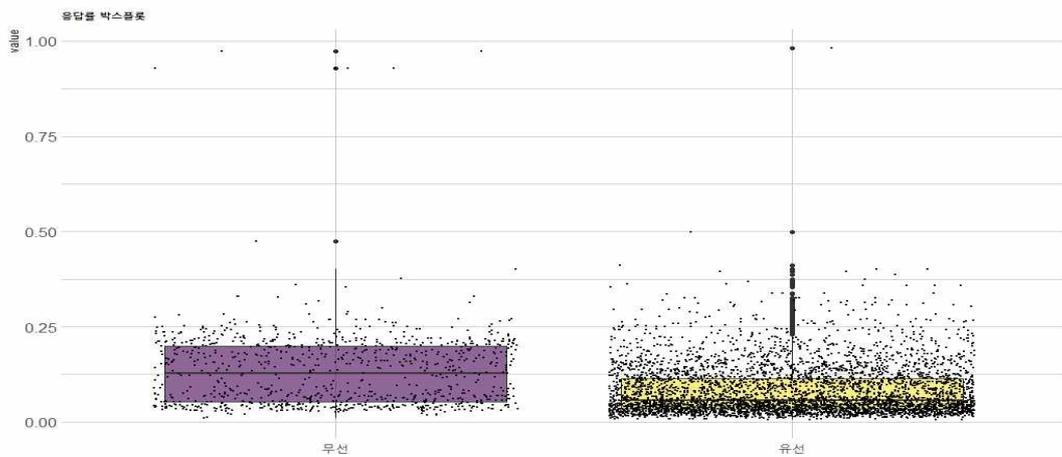
응답률: (응답 성공 사례 수)/(접촉 성공 사례수)

접촉률에서 총피조사수에서 비적격 사례수를 빼면 접촉 성공 사례 수와 접촉 성공 실패수의 합과 동일해진다. 일반적으로 할당추출하에서 비적격은 (i) 전통적 비적격과 (ii) 할당특성에 따른 비적격으로 나뉠 수 있는데, 후자는 임의로 비적격 처리한 것이기 때문에 실제 접촉률 계산에서는 포함되어야 한다. 하지만 할당특성에 따른 비적격을 분리하여 리포트하지 않기 때문에 현 보고서에서도 이를 분리하지 않고 그대로 사용하였다. 할당에 따른 비적격 사례 발생 여부 및 기타 분석은 실험 데이터 분석에서 다뤄질 예정이다.

<그림 3-1>은 접촉조사자 수를 ‘적격피조사자수’(=접촉성공수+접촉실패수)로 나눈 값들의 박스플롯이다. 유선전화번호를 사용했을 경우의 평균 접촉률은 0.43이고 중위값은 0.41이고 표준편차는 0.173이다. 한편, 무선전화번호를 사용했을 경우의 평균 접촉률은 0.37, 중위값은 0.40이고 표준편차는 0.165이다. 유선전화번호를 사용했을 경우에 접촉률이 약간 더 높은 것을 확인할 수 있는데, 이는 전화번호확인 유무로 설명이 가능한 것으로 사료된다.



<그림 3-1> 유·무선전화번호 접촉률



<그림 3-2> 유선/무선전화번호 응답률 비교

<그림 3-2>는 응답률에 대한 박스플롯 그림이다. 유선전화번호를 사용했을 경우의 응답률 평균은 0.08, 중위값은 0.06 그리고 표준편차는 0.065이다. 무선전화번호를 사용했을 경우에는 응답률 평균과 중위값 모두 0.13이고 표준편차는 0.10이다. 유선전화번호를 사용했을 경우보다 무선전화번호를 사용했을 경우에 응답률이 5%

정도 더 높은 것을 확인할 수 있다.

조사방법을 ARS와 전화면접으로 세부화 하였을 경우의 접촉률과 응답률은 <표 3-16>에 정리되어 있다. <표 3-16>은 분석의 편의상 단독조사 방법으로만 조사된 결과에 한정해서 결과를 요약하였다. ARS 방법을 사용할 경우에 접촉률은 높은편이지만 응답률은 유·무선에 상관없이 낮아지는 것을 확인할 수 있다. 무선전화면접의 경우 접촉하기는 쉽지 않지만 접촉이 되고나면 응답에 대한 성공률은 다른 조사방법 보다 높은 것을 확인할 수 있다.

<표 3-16> 조사방법별 접촉률과 응답률 (단독조사)

조사방법		평균(중위값)접촉률	평균(중위값)응답률
유선	ARS	0.49 (0.48)	0.05 (0.04)
	전화면접	0.46 (0.41)	0.14 (0.12)
무선	ARS	0.45 (0.46)	0.07 (0.06)
	전화면접	0.33 (0.33)	0.20 (0.17)

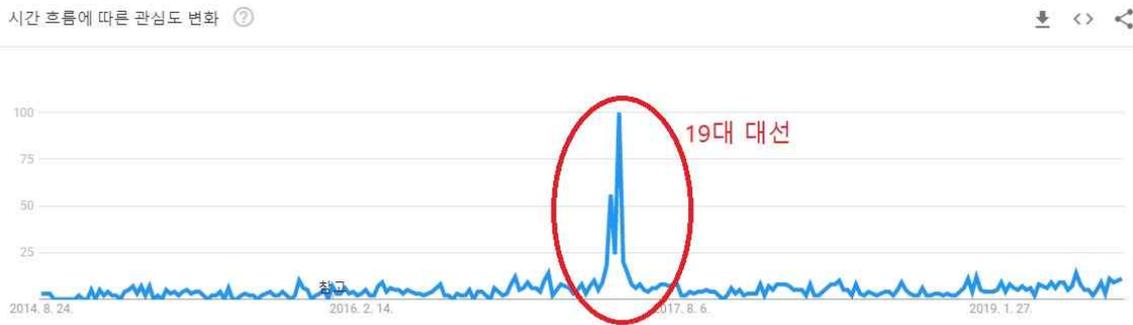
3.3 심화 분석

3.2 절에서는 여론조사 결과에 기초 분석을 실시하였다. 하지만 여론조사의 데이터의 특성상 데이터가 생성되는 과정에 대한 제어 없이 비교하는 것은 한계가 있다. 이러한 문제를 해결하기 위해서는 실험 데이터를 진행하거나 혹은 각 조사 기간 내에 유사한 조사들을 대상으로 비교 분석할 필요가 있다. 따라서 본 절에서는 조사대상을 고정한 뒤에 특정 기간에 실시된 조사 결과를 비교하는 심층 분석을 진행하였다. 조사기관들을 서로 비교하는 대신에 조사기관 내에서 조사방법에 따라서 조사값들이 어떻게 변화하는지를 살펴보았다.

1) 데이터

조사기간과 대상을 선정하기 위하여 구글 트렌드에 주요 여론조사 기관들을 키워드로 하여 구글 트렌드 키워드를 진행하였다. <그림 3-3>에 나오듯이 19대 대선 직

전에 여론조사 기관들에 대한 검색이 활발했던 것을 확인할 수 있다. 또한, 일반적으로 대통령선거는 조사가 편리하고 예측의 정확도와 높은 조사로 알려져있기 때문에 비교 분석에도 용이할 수 있다. 이러한 점을 고려해서 2017년 대통령선거 직전의 한 달(2017년 4월 2일-5월 2일) 동안 실시된 여론조사를 분석 대상으로 선정하였다. 해당기간에 중앙선거여론조사심의위원회 홈페이지에서 수집된 대통령후보 지지율 조사는 총 107개이다. 이 중에서, ARS 방법을 주요한 방법으로 사용한 여론조사 기관 중에서 조사 횟수가 가장 높은 2개업체를 선정하였고, 동일하게 전화면접을 주요한 조사 방법으로 사용하는 여론조사 기관들 중에서도 조사 횟수가 가장 높은 2개 업체를 선정하였다. 선정된 조사기관과 조사기관이 수행한 여론조사는 <표 3-17>과 <그림 3-4>에 요약되어 있다.



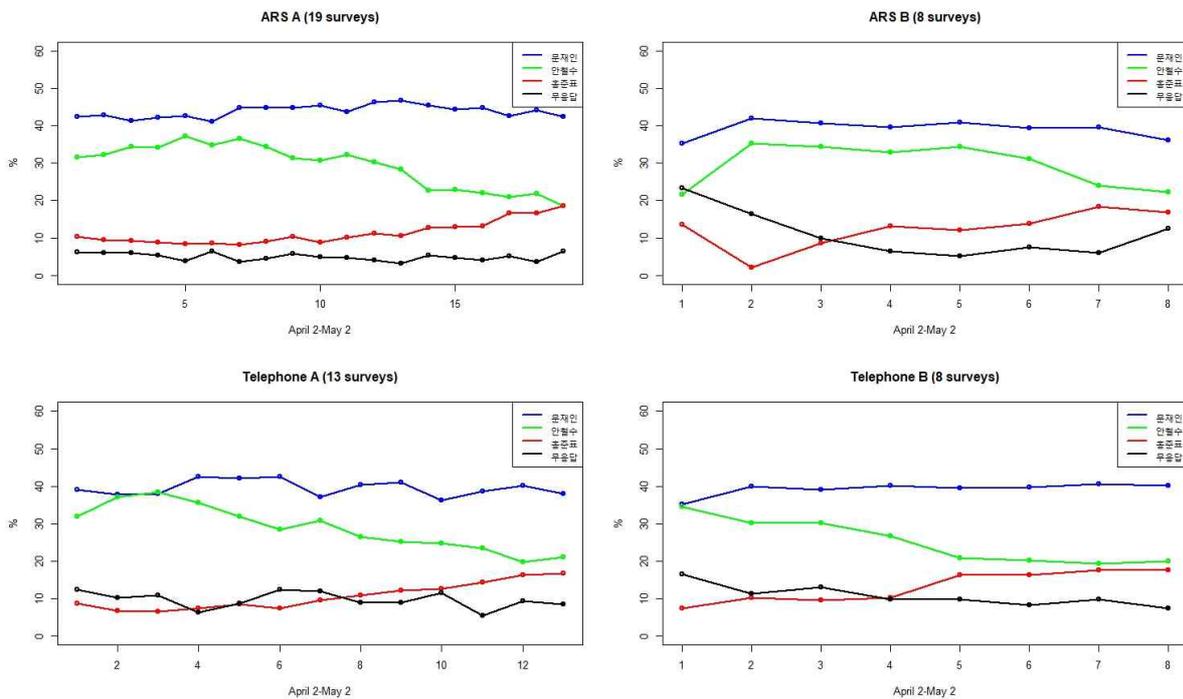
<그림 3-3> 주요 여론조사 기관들의 구글 트렌드 키워드 5년간 시계열 추이

<표 3-17> 조사방법별 주요 조사기관의 조사횟수 및 조사방법 비교

조사기관	조사횟수	주요 조사방법
ARS A	19 회	유·무선 ARS
ARS B	8 회	유·무선 ARS
TEL A	13 회	유·무선 전화면접
TEL B	8 회	유·무선 전화면접

2) ARS 기반 조사기관 사례 분석

<표 3-18>에는 조사업체 (ARS-A)가 2017년 4월에서 5월 초까지 실시한 19회의 여론 조사결과가 요약되어 있다. 해당 업체는 18번의 조사에서는 유선ARS, 무선ARS, 무선전화면접을 혼합하여 조사하였고, 유선과 무선 ARS의 두 가지를 혼합하여 1번 조사하였다.



<그림 3-4> 조사방법별 주요 여론조사 기관의 후보간 지지율 추이

<표 3-19>에는 조사업체 (ARS-B)가 실시한 8회의 여론조사 결과가 요약되어 있다. 해당 업체는 8번의 조사 중에서 7번은 유선ARS와 무선ARS를 혼합하여 사용하였고 마지막 조사의 경우에는 유선과 무선 전화면접을 혼합하여 사용하였다. 여론조사업체 ARS_B가 실시한 여론조사에서 발견할 수 있는 특이점은 두 가지로 (i) 한 가지는 유무선 혼합비가 일정하지 않다는 것이고, (ii) 다른 하나는 혼합비에 따라서 조사 결과에 차이가 큰 조사가 있다는 점이다. 즉 업체 ARS-A에 비교하여 혼합비가 일정하지 않은 편이고, 4월 1일에 시작되고 4월 3일에 완료된 두 조사에서 후보들의 조사 결과에 대한 차이가 표본오차 범위를 훨씬 넘어서 발생하였다.

<표 3-18> ARS-A 업체 조사 결과 요약

시작일	종료일	문재인	안철수	홍준표	조사방법	방법비율 (%)
04-03	04-04	42.4	31.5	10.3	유무선ARS, 무선전화	(10, 71, 19)
04-03	04-05	42.9	32.2	9.4	유무선ARS, 무선전화	(10, 71, 19)
04-05	04-05	41.3	34.5	9.2	유무선ARS, 무선전화	(10, 68, 22)
04-03	04-07	42.2	34.1	8.9	유무선ARS, 무선전화	(10, 71, 19)
04-07	04-08	42.6	37.2	8.4	유무선ARS, 무선전화	(10, 39, 51)
04-08	04-09	41.1	34.8	8.6	유무선ARS, 무선전화	(10, 69, 21)
04-10	04-12	44.8	36.5	8.1	유무선ARS, 무선전화	(10, 72, 18)
04-10	04-14	44.8	34.4	9.0	유무선ARS, 무선전화	(10, 72, 18)
04-13	04-14	44.8	31.3	10.3	유무선ARS, 무선전화	(10, 72, 18)
04-14	04-14	45.4	30.7	8.9	유무선ARS	(10, 90)
04-17	04-18	43.8	32.3	10.2	유무선ARS, 무선전화	(10, 35, 55)
04-17	04-18	46.4	30.2	11.1	유무선ARS, 무선전화	(10, 71, 19)
04-19	04-21	46.7	28.4	10.5	유무선ARS, 무선전화	(10, 70, 20)
04-24	04-25	45.5	22.6	12.8	유무선ARS, 무선전화	(10, 70, 20)
04-24	04-26	44.4	22.8	13.0	유무선ARS, 무선전화	(10, 70, 20)
04-25	04-26	44.7	22.1	13.1	유무선ARS, 무선전화	(20, 60, 20)
04-27	04-29	42.6	20.9	16.7	유무선ARS, 무선전화	(20, 60, 20)
04-28	04-29	44.1	21.8	16.6	유무선ARS, 무선전화	(10, 40, 50)
05-01	05-02	42.4	18.6	18.6	유무선ARS, 무선전화	(20, 29, 51)

<표 3-19> ARS-B 업체 조사 결과 요약

시작일	종료일	문재인	안철수	홍준표	조사방법	방법비율 (%)
04-01	04-03	35.3	21.6	13.6	유무선ARS	(45, 55)
04-01	04-03	41.9	35.3	2.2	유무선ARS	(90, 10)
04-08	04-10	40.6	34.4	8.7	유무선ARS	(49, 51)
04-15	04-16	39.7	32.8	13.1	유무선ARS	(48, 52)
04-15	04-17	41.0	34.3	12.0	유무선ARS	(44, 56)
04-22	04-24	39.4	31.1	13.9	유무선ARS	(47, 53)
04-30	05-01	39.5	23.9	18.4	유무선ARS	(44, 56)
05-01	05-02	36.1	22.3	16.9	유무선 전화면접	(37, 63)

3) 전화면접 기반 조사기관 사례 분석

<표 3-20>에는 조사업체 (TEL-A)가 실시한 13회의 여론조사 결과가 요약되어 있다. 해당 업체는 모든 조사에서 유무선 전화면접을 혼합하여 사용하였다. 조사방법과 관련하여 특이사항으로 보이는 것은 없고 혼합비가 대략 20대 80 수준에서 유지되었다.

<표 3-20> TEL-A 업체 조사 결과 요약

시작일	종료일	문재인	안철수	홍준표	조사방법	방법비율 (%)
04-04	04-04	39.1	31.8	8.6	유무선 전화면접	(18, 82)
04-07	04-08	37.7	37.0	6.7	유무선 전화면접	(24, 76)
04-11	04-12	38.0	38.3	6.5	유무선 전화면접	(14, 86)
04-15	04-16	42.6	35.6	7.3	유무선 전화면접	(18, 82)
04-18	04-19	42.0	31.8	8.5	유무선 전화면접	(19, 81)
04-18	04-20	42.6	28.4	7.5	유무선 전화면접	(16, 84)
04-18	04-21	37.0	30.8	9.5	유무선 전화면접	(12, 88)
04-24	04-25	40.4	26.4	10.8	유무선 전화면접	(20, 80)
04-25	04-25	41.0	25.2	12.2	유무선 전화면접	(18, 82)
04-26	04-26	36.2	24.7	12.6	유무선 전화면접	(20, 80)
04-29	04-30	38.7	23.4	14.4	유무선 전화면접	(16, 84)
05-01	05-02	40.2	19.7	16.2	유무선 전화면접	(18, 82)
05-01	05-02	38.0	21.0	16.8	유무선 전화면접	(18, 82)

<표 3-21>에는 조사업체 (TEL-B)가 실시한 8회의 여론조사 결과가 요약되어 있다. 조사업체 TEL-A와 동일하게 모든 조사에서 유무선 전화면접 혼합을 사용하였다. 대부분의 조사에서 혼합비율은 비슷한 수준으로 유지되었다. 또한, 전화면접 조사의 특이사항으로는 후보자들의 지지율 추이가 ARS 조사보다는 조사업체간 차이가 적다는 것을 확인할 수 있다.

<표 3-21> TEL-B 업체 조사 결과 요약

시작일	종료일	문재인	안철수	홍준표	조사방법	방법비율 (%)
04-07	04-08	35.2	34.5	7.4	유무선 전화면접	(14, 86)
04-18	04-19	40.0	30.1	10.2	유무선 전화면접	(17, 83)
04-21	04-22	39.1	30.1	9.5	유무선 전화면접	(21, 79)
04-26	04-26	40.1	26.8	10.2	유무선 전화면접	(21, 79)
04-29	04-30	39.4	20.8	16.2	유무선 전화면접	(20, 80)
04-30	05-01	39.7	20.2	16.4	유무선 전화면접	(21, 79)
05-01	05-02	40.6	19.3	17.7	유무선 전화면접	(20, 80)
05-01	05-02	40.2	19.9	17.7	유무선 전화면접	(21, 79)

4) 심층분석 결과 요약

여론조사는 조사기관, 조사방법, 가중치 조정 방법에 따라서 결과가 많이 달라질 수 있다. 이러한 점을 고려해서 본 절에서는 ARS와 전화면접을 주요한 조사 방법으로 진행하는 업체를 선정하 뒤 해당 업체의 지난 19대 대통령선거 여론조사의 결과를 요약하였다. 분석 결과를 중심으로 몇 가지 주요한 특이사항을 정리하면 다음과 같다.

- (i) 조사방법에 상관없이 유사 혹은 동일한 시점에서 얻어진 조사결과의 변동폭이 표본오차를 감안했을 때 작은 점과 표본오차가 응답표본크기에 의해서만 결정되는 점을 고려했을 때, 가중치 보정과 그에 따른 표본오차 계산이 제대로 이루어지고 있지 않고 있는 것으로 사료된다.
- (ii) ARS조사에서 조사업체간의 차이가 전화면접에서의 조사업체간 차이보다 큰 것을 앞 선 심층 분석에서 확인할 수 있다. 심층 분석에 사용된 ARS 조사업체들의 유무선 혼합비율이 상이하기 때문에 정확한 비교는 어렵다. 하지만 낮은 응답률 및 임의적인 혼합비율 등을 고려할 때, 다른 조사보다는 ARS 조사에서 모드 편차의 (방법론에 따른 편차)가 더 클 요인이 있는 것을 판단된다. 혼합비에 대한 적절한 검토 및 검증과정이 필요할 것으로 생각 된다.

(iii) 전체적으로 ARS를 주요한 조사 방법으로 사용하는 경우 유의미하게 무응답 비중이 낮다. 일반적으로 투표일에 가까워질수록 무당층이 줄어드는 점을 고려할 때, ARS/전화면접에서의 무응답(및 할당) 처리 방법이 다르다는 것을 짐작할 수 있으며 이 부분은 실험데이터 분석에서 추가적으로 논의될 예정이다.

3.4 소결

본 절에서는 선거 여론조사를 다차원적으로 비교 분석하면서 현황 및 실태를 파악하고 관련 내용을 정리하였다. 분석 내용을 근거로 세 가지 측면(조사비용 현실화, 조사방법 과학화, 사후 검정 과정 체계화)에서 선거 여론조사를 평가하고 발전 방향을 모색할 수 있을 것으로 사료된다.

1) 조사비용 현실화

현재 선거 여론조사의 품질을 높이기 위해서는 우선적으로 조사비용의 현실화가 이뤄져야할 것으로 사료된다. (i) 대부분의 선거 여론조사가 2일 이내로 집계되고, (ii) 조사단가가 저렴한 대신 선택 편향(selection bias)에 대하여 상대적으로 취약한 ARS 기반 조사가 전화면접에 비하여 훨씬 더 많이 사용되고 있는 이유는 1차적으로 비용 문제에 기인할 것으로 추측된다. 조사품질 향상을 위해서는 조사 비용을 현실화할 필요가 있으며, 이를 위해서는 현행 선거 여론조사의 시장 규모 및 조사비용에 대한 조사와 후속 연구가 필요할 것으로 생각된다. 이때, 자료 수집에 드는 비용뿐만 아니라 체계적인 관리 및 분석까지를 비용으로 고려해야할 것이다.

2) 조사방법 과학화

현재 거의 모든 조사는 (i) 할당추출에 근거한 비확률표본을 사용하고 있으며, 따라서 (ii) 통계적인 표본오차를 계산하지 못하고 있고, (iii) 조사방법의 특성

(mode effect) 고려 없이 혼합 조사가 실시되고 있다. 그리고 (iv) 여론조사와 예측조사가 전혀 분리되어 있지 않으며, (v) 획일적인 가중화 작업이 이뤄지고 있다. 이러한 문제는 부분적으로 조사비용에 기인한 측면도 있지만, 기본적으로는 과학적인 조사방법에 대한 지속적인 고민과 투자가 없기 때문인 것으로 사료된다. 일부 조사기관(한국조사협회)의 경우 과학적인 조사방법을 지향하고 있는 것으로 파악되지만, 전체 선거여론조사 시장의 규모와 영향력을 고려했을 때는 아직 미약한 것으로 판단된다. 조사방법의 과학화가 모든 조사기관으로 확대되어 꾸준히 관리될 때, 신뢰성 있는 조사결과가 생성될 수 있을 것으로 기대한다.

3) 사후 검증 과정 체계화

선거철마다 조사업체들이 난립하고 검증되지 않은 방법들이 마구잡이로 사용되는 이유 중의 하나로 사후 검증 과정 체계의 부재를 들 수 있다. 우선적으로 조사원자료가 일정 기간 (예: 5년) 보관되어야 할 것이며, 또한 최종 추정값에 사용된 일련의 모드 효과 보정 및 가중화 방법이 기록되어야 할 것이다. 이를 통해서, 비체계적이고 과학적인 방법을 사용하는 조사기관들이 지양될 수 있도록 해야 할 것이다.

제 4 장 실험연구를 통한 표본추출틀 및 조사모드의 체계적 차이 검토

4.1 실험설계 개요 및 조건

본 연구를 위한 실험연구는 현행 선거여론조사의 이해와 현상파악을 주요 목적으로 하였다. 주어진 연구과제의 예산 하에서 표본추출틀과 조사모드에 따른 선거여론조사의 표본선택, 조사운영, 응답분류 및 조사결과를 비교하고자 하였다.³⁾ 표본추출틀로는 유선전화번호 RDD(유선RDD), 무선전화번호 RDD(무선RDD), 그리고 가상번호를 고려하였다. 조사모드로는 전화면접조사와 ARS조사의 두 가지 조사방식을 고려하였다. <표 4-1>은 본 연구를 위해 고려한 표본추출틀과 조사모드에 의해 결정되는 총 4가지의 실험조건과 목표응답수를 정리하고 있다. 유무선 RDD 혼합표본과 가상번호와 유선 RDD 혼합표본에 대해 각각 전화면접조사와 ARS전화조사를 수행하는 것이다. 실험조건별로 각각 500개의 응답을 목표로 설정하였다.

4.2 실험연구 진행

실험연구는 전화면접조사와 ARS조사를 전문으로 수행하는 조사기관 각 1곳을 선정하여 조사를 의뢰하였다. 조사는 9월30일에서 10월2일까지 총 3일간 <표 4-2>에 요약된 일정으로 진행하였다. <표 4-3>은 실험연구를 위한 설문문항을 요약하고 있고 부록 A-1과 A-2에 각각 전화면접조사와 ARS조사에 사용된 문항을 포함하고 있다.

3) 조사모드에 따른 표본응답자에 대한 선택편향(selection bias)과 응답값에 대한 모드효과(mode effect)를 적절히 평가하기 위해서는 한 조사회사를 통해 전화면접과 자동응답을 수행하는 것이 바람직할 수 있다. 하지만, 대규모 여론조사 회사들이 주로 회원사로 있는 한국조사협회는 2014년 'ARS조사 관련 한국조사협회 회원사의 행동규범'의 공표를 통해 향후 ARS 조사를 수행하지 않기로 하고, 선거여론조사에 대해서는 전화면접조사만 사용하고 있다. 반면, 주로 중소기업체들로 이루어진 '한국정치조사협회'의 회원사들만 ARS 조사를 수행하는 것으로 파악된다. 따라서 본 연구에서 실험설계는 현행 선거여론조사의 이해와 현상파악으로 연구목표를 축소하였다.

<표 4-1> 실험조건별 목표 응답수

(단위 : 명)

실험조건		조사모드 (측정)		
		전화면접조사(M_1)	ARS조사(M_2)	총합
표본추출틀 (조사대상 구성)	유무선 (F_1)	500	500	1,000
	가상번호+ 유선 (F_2)	500	500	1,000
	총합	1,000	1,000	2,000

<표 4-2> 실험연구를 위한 선거여론조사 업무별 일정

업무	일정	비고
조사회사 발주	9월16일~9월20일	
가상번호요청	9월20일	조사개시 10일전 신청
조사 실시 신고	9월27일	조사개시 3일전 신고
조사 일정	9월30일~10월2일 (3일간)	
분석결과 수취	10월4일	

<표 4-3> 설문문항

구분	질문	
기본질문	연령	연세는 만으로 어떻게 되십니까?
	성별	성별은 무엇입니까?
	지역	주소지 기준으로 어느 시도에 살고 계십니까?
여론문항	대통령 국정수행 최근 이슈	현재 문재인 대통령의 국정수행에 대해 어떻게 평가하십니까? 신임 조국 법무부 장관의 장관직 수행에 대해 어떻게 생각하십니까?
	성향 문항	정치관심도 투표의향 정치성향 여론조사 효용성 정당지지도
응답자특성	직업	직업은 무엇입니까?
	학력	학교를 어디까지 마치셨습니까? 중퇴는 졸업에 포함되지 않습니다.

4.3 실험연구 비교분석

1) 최종적 실험조건 비교

실험진행 후 확인한 결과, ARS조사에서 원래 조사계획과는 달리 유선RDD와 가상번호만을 사용하여 조사가 진행되었다. 이에 유무선 RDD 혼합번호를 이용한 조사를 추가로 요청하여 10월10일~11일 양일간에 걸쳐 조사하였다. <표 4-4>는 최종 실험조건을 정리하고 있다. 총 5개의 실험조건이 고려되었고 각각 500여명의 실질적 응답자수가 달성되었다. 이후 주요문항별 조사결과의 비교는 조사회사에서 각 조건별로 산출한 표본가중치를 적용한 추정값을 이용하였다.

<표 4-4> 최종 실험조건 구성 비교

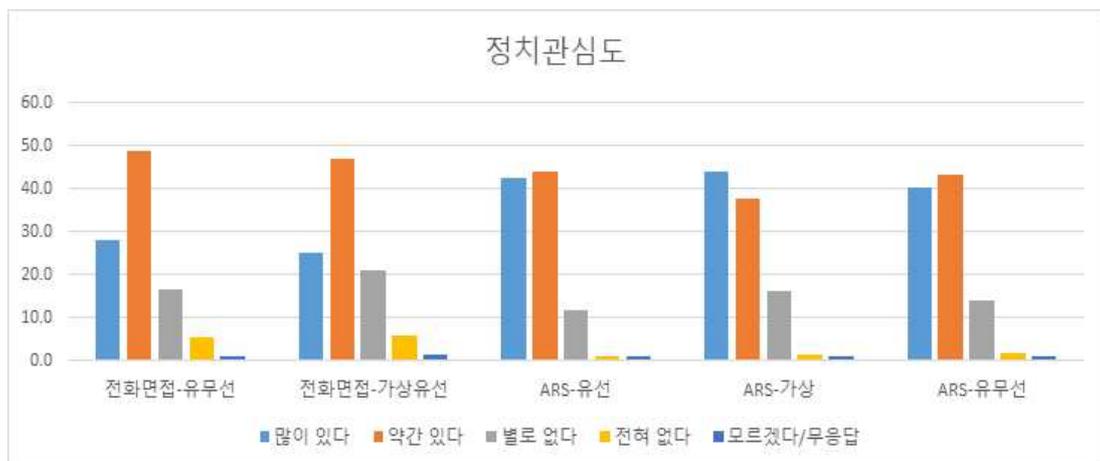
조건	조사모드	표본응답구성	조사일정
1. 전화면접-유무선	전화면접조사	502명 (=무선400+유선102)	'19. 9.30~10.2(3일)
2. 전화면접-가상유선	전화면접조사	502명 (=가상408+유선 94)	'19. 9.30~10.2(3일)
3. ARS-유선	ARS조사	538명 (=유선538)	'19. 9.30~10.2(3일)
4. ARS-가상	ARS조사	513명 (=가상513)	'19. 9.30~10.2(3일)
5. ARS-유무선	ARS조사	518명 (=무선407+유선111)	'19.10.10~10.11(2일)

2) 정치관심도

<표 4-5>와 <그림 4-1>는 실험조건별 정치관심도의 응답구성을 비교하고 있다. 정치관심도에 대해서는 표본(응답)구정보다는 조사모드 간 차이가 두드러지고 있는데, 예로, ARS조사에서 정치관심도가 ‘많이 있다’라는 비율이 40.2~44.0%로 높은 반면, 전화면접조사에서는 ‘약간 있다’의 비율이 46.8~48.7%로 높았다. 또한 정치관심도가 ‘별로 없다’와 ‘전혀 없다’의 비율은 ARS조사에서 보다 전화면접조사에서 많이 나타났다.

<표 4-5> 실험조건별 정치관심도 비교

전화모드	표본(응답)구성	정치관심도				
		많이 있다	약간 있다	별로 없다	전혀 없다	모르겠다/무응답
전화면접 조사	유무선	28.1	48.7	16.7	5.4	1.1
	가상번호+유선	25.2	46.8	20.9	5.7	1.4
ARS 조사	유선	42.6	44.0	11.6	1.0	0.8
	가상번호	44.0	37.8	16.0	1.2	1.0
	유무선	40.2	43.4	14.0	1.6	0.8



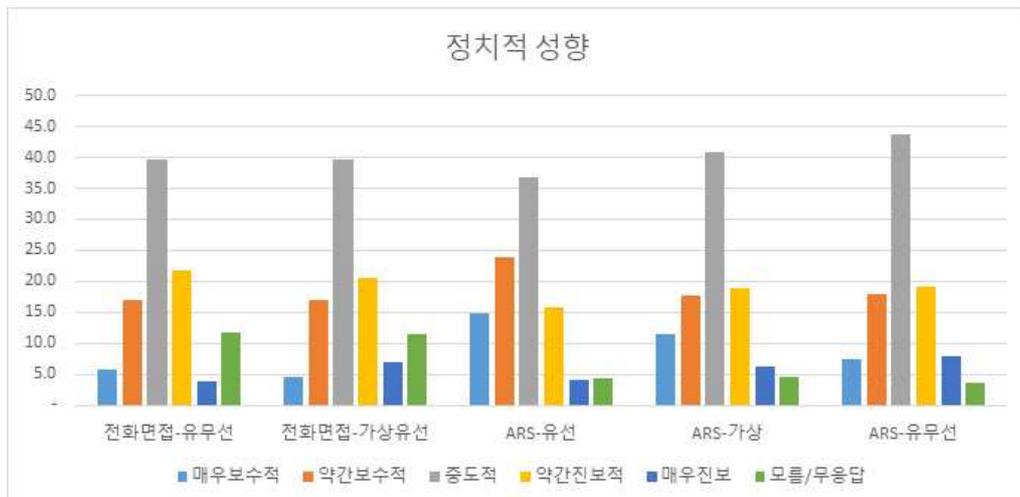
<그림 4-1> 실험조건별 정치관심도 비교

3) 정치적 성향

<표 4-6>과 <그림 4-2>는 실험조건별로 정치적 성향에 대한 응답구성을 비교하고 있다. ARS조사에서 유선표본의 응답자들은 가상번호나 유무선 표본에 비해 ‘매우보수’와 ‘약간보수’의 비율이 두드러지게 높았다. ‘매우 보수적’의 비율은 ARS조사가 전화면접조사에 비해 높게 나타났다. 또한 ‘모르겠다/무응답’은 조사모드 내 표본(응답)구성별 차이는 없지만 전화면접조사에서 11.4~11.8%로 ARS조사의 3.6~4.6%에 비해 매우 높게 나타났다.

<표 4-6> 실험조건별 정치적 성향 비교

전화모드	표본(응답)구성	정치적 성향					
		매우 보수적	약간 보수적	중도적	약간 진보적	매우 진보적	모르겠다 / 무응답
전화면접 조사	유무선	5.9	17.1	39.6	21.8	3.9	11.8
	가상번호+유선	4.6	16.9	39.7	20.7	6.9	11.4
ARS 조사	유선	14.8	24.0	36.8	15.8	4.2	4.4
	가상번호	11.6	17.8	40.8	19.0	6.2	4.6
	유무선	7.4	18.0	43.8	19.2	8.0	3.6



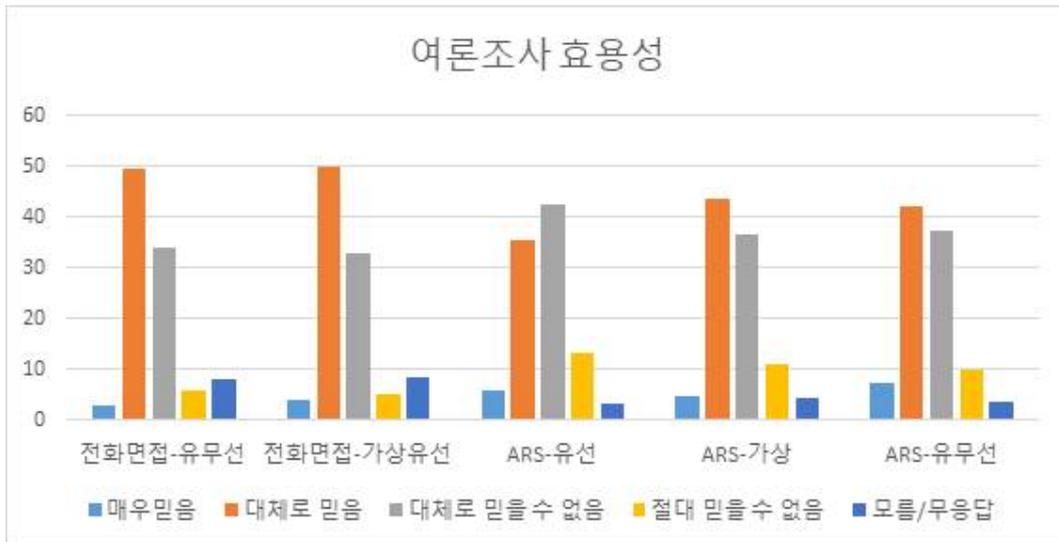
<그림 4-2> 실험조건별 정치적 성향 비교

4) 여론조사 효용성

<표 4-7>과 <그림 4-3>은 실험조건별 여론조사 효용성에 대한 의견분포를 정리하고 있다. ARS조사는 ‘여론조사에 대해 어떻게 생각하십니까?’라는 질문에 ‘매우 믿을 만하다’가 4.6~7.2%, ‘대체로 믿을 수 없다’는 36.4~42.4%, ‘절대 믿을 수 없다’는 9.8~13.0%로 극단적 의견이 전화면접조사에 비해 상대적으로 많은 반면, 전화면접조사에서는 ‘대체로 믿을 만하다’는 49.6~50.0%로 ARS조사에서 보다 다소 높게 나타났다. 본 질문에 대한 ‘모르겠다/무응답’의 경우에도 ARS에서 보다 전화면접조사에서 더 높게 나타났다.

<표 4-7> 실험조건별 여론조사 효용성 비교

조사모드	표본(응답)구성	여론조사 효용성				
		매우 믿을만하다	대체로 믿을만하다	대체로 믿을수 없다	절대 믿을수 없다	모르겠다/무응답
전화면접조사	유무선	2.9	49.6	33.9	5.8	7.8
	가상번호+유선	3.8	50.0	32.9	5.1	8.2
ARS조사	유선	5.8	35.6	42.4	13.0	3.2
	가상번호	4.6	43.8	36.4	10.8	4.4
	유무선	7.2	42.0	37.4	9.8	3.6



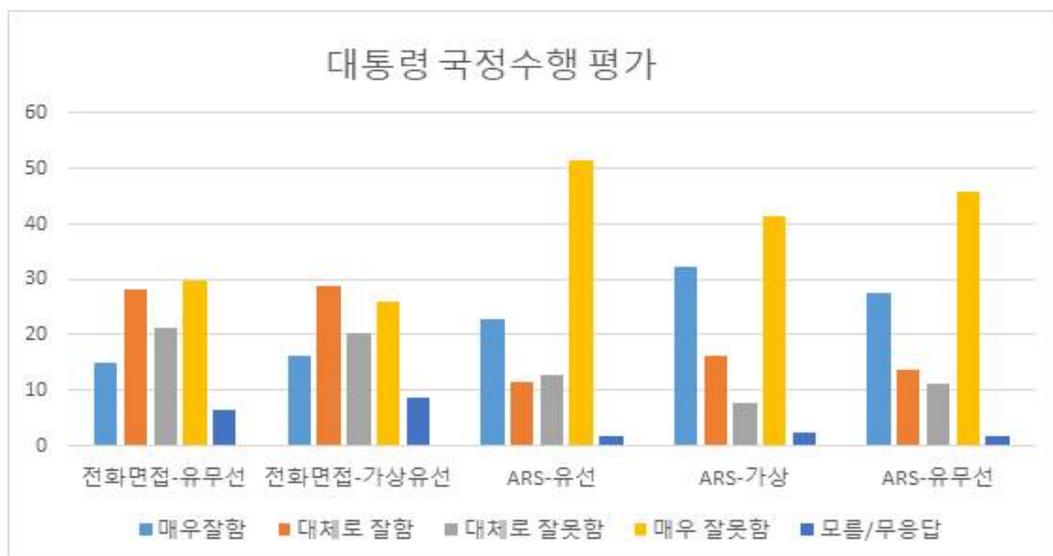
<그림 4-3> 실험조건별 여론조사 효용성 비교

5) 대통령 국정수행 평가

<표 4-8>과 <그림 4-4>는 실험조건별 대통령 국정수행 평가에 대한 의견분포를 비교하고 있다. 응답구성은 조사모드별로 그 특징이 나뉘며 조사모드 내 표본구성 간에는 큰 차이를 보이지 않는다. 전화면접조사는 ‘대체로 잘함’과 ‘대체로 잘못함’의 중간적 의견이 많은 반면, ARS조사는 ‘매우 잘함’과 ‘매우 잘못함’의 의견이 압도적으로 많음을 알 수 있다. ‘모르겠다/무응답’의 경우에도 ARS조사에서 보다 전화면접조사에서 더 높은 비중을 보였다.

<표 4-8> 실험조건별 대통령 국정수행 평가 비교

조사모드	표본(응답)구성	대통령 국정수행 평가				
		매우 잘하고 있다	대체로 잘하고 있다	대체로 잘못하고 있다	매우 잘못하고 있다	모르겠다/무응답
전화면접 조사	유무선	14.8	28.1	21.3	29.6	6.4
	가상번호+유선	16.3	28.8	20.3	25.9	8.7
ARS 조사	유선	22.8	11.4	12.6	51.4	1.8
	가상번호	32.2	16.2	7.8	41.4	2.4
	유무선	27.6	13.6	11.2	45.8	1.8



<그림 4-4> 대통령 국정수행 평가 비교

6) 실험조건별 산출률 비교

<표 4-9>는 조사평가를 위한 표본전화번호에 대한 조사결과분류를 정리하고 있다. 먼저, 응답완료와 거절/중단은 조사적격으로 구분된다. 통화중 혹은 부재중(접촉 안됨)은 적격여부를 구분할 수 없는 경우들이다. 결번, 할당초과, 대상자 아님, 사업체번호, 팩스번호 등은 부적격이고 미사용 역시 부적격으로 구분된다. 우리나라 선거여론조사는 특정한 할당특성(성·연령·지역)에 따라 응답자 구성비가 모집단 구성비와 같아질 수 있도록 조사를 수행하므로 할당초과는 미사용과 더불어 부적격으로 구분된다. 할당추출이 아닌 일반적 표본조사에서는 할당규칙을 주지 않으므로 앞서의 할당초과와 같은 상황은 발생하지 않는다. 또한 미사용도 응답자들에 대한 할당비율을 조사기간 중 이미 맞추어진 상황이라 더 이상 준비해둔 표본전화번호를 사용하지 않았음을 나타낸다. 한 가지 눈여겨 볼 것은 ARS조사는 기계에 의해 조사가 진행되므로 전화면접조사와는 달리 거절, 사업체번호, 팩스번호 등을 구분할 수 없다는 단점을 갖는다.

<표 4-9> 조사평가를 위한 조사결과분류

구분자	분류	조사적격성	조사모드	
			전화면접조사	ARS조사
<i>I</i>	조사완료	적격	응답완료	응답완료
<i>R</i>	거절/중단	적격	중도이탈 거절	중도이탈
<i>U</i>	적격불명	적격불명	통화중 부재중(접촉 안 됨)	통화중 접촉안됨
<i>NE</i>	부적격	부적격	결번 할당초과 대상자 아님 사업체번호 팩스번호 미사용	결번 할당초과 대상자 아님

< 표 4-10>와 <표 4-11>은 실험조건별 번호종류에 따른 조사결과를 표본전화 번호수와 상대크기(퍼센트)로 각각 보여주고 있다. 예로, 전화면접조사 1은 유선 RDD 번호 3,765개와 무선 RDD 번호 8,000개로 구성된 총 11,765개의 번호를 표집 하였다. 이중 최종 응답수는 유선 RDD 번호 102개 (2.7%)와 무선 RDD 번호 400개 (5.0%)인 총 502개의 응답을 얻었다. 유선 RDD 번호인 경우, 거절/중도이탈은 503 개(13.4%), 통화중/부재중으로 인한 비접촉은 1,038개(27.6%)이었고, 나머지 2,122 개는 비적격으로 결번이 1,719개(45.7%)로 제일 많고, 미사용 236개 (6.3%), 할당 초과 127개 (3.4%), 대상자 아님이 18개 (0.5%), 사업체팩스 22개 (0.6%)이었다. 무선 RDD 번호인 경우, 거절/중도이탈은 1,608개 (20.1%), 통화/부재로 인한 비접촉 은 3,842개 (48.0%)이었고, 나머지 2,150개는 비적격으로 결번이 1,589개 (19.9%)로 제일 많고, 미사용 0개, 할당초과 452개 (5.7%), 대상자 아님이 96개 (1.2%), 사 업체팩스 13개 (0.3%)이었다.

유선 RDD의 결번 비중은 실험조건별로 45.7%~57.9%로 높은 편인데 이는 각 조사회사에서 확보한 국번에 사용가능한 1만개 번호의 임의번호를 생성하는 RDD 번 호 프레임이 그 만큼 효율적이 않다는 것을 보여준다. 무선 RDD의 결번 비중은 전화 면접조사 1에서는 19.9%이고 ARS 조사에서는 22.3%로 유선 RDD의 결번 비중에 비 해 낮은 수준이었다. 거절/중도이탈의 비중은 전화면접조사에 비해 ARS조사에서 높 게 나타나는데, 유선 RDD 번호에 대해서는 전화면접조사에서 13.4%~17.9%인 반면 ARS조사에서는 23.3~24.4%이고, 무선 RDD와 가상번호에서는 전화면접조사는 20.1%와 14.3%인데 반해 ARS 조사는 45.5%와 50.3%로 매우 높게 나타났다. 전화 면접조사 2의 가상번호에 대한 미사용 비중이 33.8%로 매우 높는데, 이는 본 연구를 위한 조사에서는 기존 조사에서 보다는 더욱 많은 노력을 기울인데 따른 결과로 판단 된다.

전체 표본수 대비 응답수, 즉 응답완료률은 전화면접조사는 2.7~6.8%이고 ARS 조사는 0.6~3.6%로 매우 낮다. 이 중 가상번호는 전화면접조사에서 6.8%와 ARS조 사에서 3.6%로 다른 표본전화번호의 종류에 비해 상대적으로 높게 나타남을 알 수 있다.

<표 4-12>는 실험조건별로 조사결과에 대한 다양한 산출물을 요약하고 있다.

사용된 산출률은 <표 4-13>의 정의를 사용하였다. 먼저, 중앙선거여론조사심의위원회에서 현재 사용하고 있는 접촉률은 추정된 적격 대상 전화번호 중 접촉비이고 응답률은 확인된 적격 대상 전화번호 중 응답비로 정의한다. 더불어, 표본조사에서 일반적으로 참고되는 미국여론조사학회(American Association for Public Opinion Research, AAPOR)의 산출률은 다음과 같다. 접촉률은 비적격을 제외한 대상 중 접촉률이고, 적격률은 적격성이 확인된 대상 표본번호 중 적격 번호 비율이며, 협조율은 확인된 적격 표본전화번호 중 응답 비율이고, 응답률은 추정된 적격 대상 중 응답비로 각각 정의된다. 여심위의 응답률은 AAPOR의 협조율에 해당하며, 여심위의 접촉률과 응답률을 곱하면 AAPOR의 응답률과 같게 된다.

<표 4-14>는 실험연구가 수행되었던 즘에 실시된 전화면접조사와 ARS조사로 중앙선거여론조사심의위원회 홈페이지에 등록된 유선과 무선RDD 부분에 대한 중앙선거여론조사심의위원회 기준의 접촉률과 응답률을 정리하고 있다. 실험연구에서 실험조건별로 얻어진 접촉률과 응답률은 여심위 홈페이지에 등록된 타 조사의 그것들에 비해 월등히 높음을 알 수 있다. 예로, 유선 RDD 번호에 대한 전화면접조사와 ARS조사의 접촉률의 경우, 실험연구에서는 70.6~74.2%인 반면 한국갤럽(등록기준: 10월 10일, 조사일: 10월8일/10월10일)과 리얼미터(등록기준: 10월11일, 조사일: 10월7일~8일/10월10일~11일)의 조사에서는 29.8%와 33.8%로 매우 낮게 보고되고 있다. 본 연구를 위해 수행된 조사에서는 조사기관들이 평소에 수행하던 조사에서보다 매우 많은 노력(예, 재접촉 등)을 하였던 것으로 보인다.

<표 4-10> 실험조건별 응답분류

(단위: 개)

응답구분	전화면접조사 1			전화면접조사 2			ARS 조사			ARS 조사 (추가)		
	유선 RDD	무선 RDD	전체	유선 RDD	가상 번호	전체	유선 RDD	가상 번호	전체	유선 RDD	무선 RDD	전체
전체 사용규모 (T)	3,765	8,000	11,765	3,000	5,985	8,985	87,390	14,181	102,390	19,810	34,491	54,301
응답완료(I)	102	400	502	94	408	502	538	513	1,051	111	407	518
거절/중도이탈(R)	503	1,608	2,111	450	858	1,308	20,377	7,133	27,510	4,827	15,667	20,494
비적격-결번 (NE)	1,719	1,589	3,308	1,537	90	1,627	50,564	4	50,568	10,970	7,687	18,657
비적격-미사용 (NE)	236	-	236	12	2,022	2,034		819	819		-	-
비적격-할당초과(NE)	127	452	579	119	46	165	1,796	-	1,796	197	481	678
비적격-대상자아님(NE)	18	96	114	12	8	20	38		38	17	49	66
비적격-사업체팩스(NE)	22	13	35	6	22	28	-	-	-	-	-	-
비접촉-통화/부재 (U)	1,038	3,842	4,880	770	2,531	3,301	14,077	6,531	20,608	3,688	10,200	13,888

<표 4-11> 실험조건별 응답분류

(단위: %)

응답구분	전화면접조사 1			전화면접조사 2			ARS 조사			ARS 조사 (추가)		
	유선 RDD	무선 RDD	전체	유선 RDD	가상 번호	전체	유선 RDD	가상 번호	전체	유선 RDD	무선 RDD	전체
전체 사용규모 (T)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
응답완료(I)	2.7	5.0	4.3	3.1	6.8	5.6	0.6	3.6	1.0	0.6	1.2	1.0
거절/중도이탈(R)	13.4	20.1	17.9	15.0	14.3	14.6	23.3	50.3	26.9	24.4	45.4	37.7
비적격-결번 (NE)	45.7	19.9	28.1	51.2	1.5	18.1	57.9	0.0	49.4	55.4	22.3	34.4
비적격-미사용 (NE)	6.3	-	2.0	0.4	33.8	22.6	-	5.8	0.8	-	-	-
비적격-할당초과(NE)	3.4	5.7	4.9	4.0	0.8	1.8	2.1	-	1.8	1.0	1.4	1.2
비적격-대상자아님(NE)	0.5	1.2	1.0	0.4	0.1	0.2	0.0	-	0.0	0.1	0.1	0.1
비적격-사업체팩스(NE)	0.6	0.2	0.3	0.2	0.4	0.3	-	-	-	-	-	-
비접촉-통화/부재 (U)	27.6	48.0	41.5	25.7	42.3	36.7	16.1	46.1	20.1	18.6	29.6	25.6

<표 4-12> 실험조건별 산출률

기준	산출률	전화면접조사 1			전화면접조사 2			ARS 조사			ARS 조사 (추가)		
		유선 RDD	무선 RDD	전체	유선 RDD	가상	전체	유선 RDD	가상	전체	유선 RDD	무선 RDD	전체
AAPOR	접촉률	36.8%	34.3%	34.9%	41.4%	33.3%	35.4%	59.8%	53.9%	58.1%	57.2%	61.2%	60.2%
	적격률(e)	24.3%	48.3%	39.3%	24.5%	88.4%	49.6%	28.5%	99.9%	35.3%	30.6%	66.2%	52.0%
	협조율	16.9%	19.9%	19.2%	17.3%	32.2%	27.7%	2.6%	6.7%	3.7%	2.2%	2.5%	2.5%
	응답률	11.9%	10.4%	11.1%	12.8%	11.6%	14.6%	2.2%	3.6%	2.9%	1.8%	1.8%	1.8%
여심위	접촉률	70.6%	52.0%	57.7%	74.2%	36.1%	52.5%	83.9%	53.9%	79.7%	81.4%	70.4%	74.4%
	응답률	16.9%	19.9%	19.2%	17.3%	32.2%	27.7%	2.6%	6.7%	3.7%	2.2%	2.5%	2.5%

<표 4-13> 산출률 정의

기준	종류	정의	설명
미국 조사연구 학회 (AAPOR)	접촉률	$CON = \frac{I+R}{I+R+U}$	비적격을 제외한 대상 중 접촉률
	적격률	$e = \frac{I+R}{I+R+NE}$	적격성이 확인된 대상 중 적격 표본개체비
	협조율	$CR = \frac{I}{I+R}$	(확인된) 적격 개체 중 응답비
	응답률	$RR = \frac{I}{I+R+eU}$	(추정된) 적격 대상 중 응답비
중앙선거여 론조사심의 위원회	접촉률	$CON^* = \frac{I+R}{I+R+eU}$	(추정된) 적격 대상 중 접촉비
	응답률	$RR^* = \frac{I}{I+R}$	(확인된) 적격 대상 중 응답비

<표 4-14> 여심위에 등록된 최근의 접촉률 및 응답률 사례

조사회사(등록일)		한국갤럽(10월10일 등록)		리얼미터(10월11일 등록)	
조사 조건		전화-유선RDD	전화-무선RDD	ARS-유선RDD	ARS-무선RDD
산출률	접촉률	29.8%	37.4%	33.8%	23.9%
	응답률	13.8%	18.3%	3.0%	6.0%

7) 실험조건별 응답성향 비교

무응답으로 인한 모수추정에 대한 영향력은 무응답오차(nonresponse error) 혹은 무응답편향(nonresponse bias)의 형태로 평가할 수 있다. 이는 무응답 현상에 대해 통계학적 이론을 접목하여 고려될 수 있다. 만약 특정 개체의 응답은 조사요청 시 확률적으로 결정되고 조사변수 y_i 와는 무관하게 개체의 보조적 특성으로 설명될 수 있다고 가정하자. 예로, I_i 는 개인 i 의 표본(포함)지시자, R_i 은 응답지시자, x_i 는 보조특성이라면 해당 개인이 갖는 응답확률인 응답성향점수(response propensity score)는 다음과 같이 정의할 수 있다 (Rosenbaum과 Rubin, 1983).

$$\phi(x_i) = \Pr(R_i = 1 | I_i = 1, x_i) \quad (4-1)$$

응답여부에 관계없이 모든 개체의 보조특성 x_i 을 알 수 있다면 로지스틱회귀(logistic regression)모형이나 프로빗(probit)모형 등을 이용하여 응답성향점수를 산출할 수 있다. 무선 RDD 번호를 제외하고 각 통신사로부터 제공받게 되는 가상번호나 유선 RDD 번호는 각각 지역·성·연령의 정보나 국번으로 유추되는 지역정보가 응답여부에 상관없이 표본전화번호에 대해 주어지게 된다. 따라서 제한적이거나 마 이러한 정보를 활용한다면 앞서의 응답성향모형을 적절히 적합할 수 있게 된다.

<표 4-15>는 전화면접조사에서 가상번호에 대해 로지스틱 응답성향모형을 적합한 결과를 정리하고 있다. 성별 정보를 제외하고 권역(서울, 경기·인천(경인), 충청, 전라, 경북, 경남, 강원·제주), 연령대(20대,30대,40대,50대,60대,70대), 통신사(SKT, KT, LG U+)는 모두 통계적으로 유의한 것으로 나타났다. 권역별로는 서울에 비해 충청이 낮은 응답성향을 갖고 강원·제주는 높은 응답성향을 보여준다. 또한 연령이 높을수록 응답성향은 높아지며, 통신사 간에도 차이를 보이고 있는데 SKT에 비해 LG U+와 KT가 각각 높은 응답성향을 갖는다.

<표 4-15> 전화면접조사 가상번호 응답성향모형 적합결과

예측변수	수준	자유도	카이제곱	유의확률	추정값
절편					-1.120
권역	(서울)	6	13.569	0.035	
	경인				-0.050
	충청				-0.649
	전라				-0.118
	경북				0.103
	경남				-0.019
	강원제주				0.377
연령	(20대)	5	13.178	0.022	
	30대				0.073
	40대				0.324
	50대				0.366
	60대				0.633
	70대이상				0.531
통신사	(SKT)	2	6.7763	0.0338	
	KT				0.349
	LG U+				0.273

<표 4-16>는 ARS조사에서 가상번호에 대해 로지스틱 응답성향모형을 적합한 결과를 정리하고 있다. 연령과 통신사의 두 변수는 모두 통계적으로 유의하며, 성별은 전화면접조사와는 달리 연령과 권역 간 상호작용(interaction term)의 형태로 통계적으로 유의한 것으로 나타났다. 권역은 전화면접조사와는 달리 전라권이 낮은 응답성향을 보인다. 연령대가 높을수록 응답성향은 높아지고 통신사에 대해서는 SKT에 비해 LG U+와 KT가 각각 높은 응답성향을 갖는다.

<표 4-16> ARS조사 가상번호 응답성향모형 적합결과

예측변수	수준	자유도	카이제곱	유의확률	추정값
절편					-4.058
권역	(서울)	6	3.3734	0.7607	
	경인				0.122
	충청				0.108
	전라				-0.027
	경북				0.355
	경남				0.105
강원제주	0.321				
연령	(20대)	5	121.339	<.0001	
	30대				0.993
	40대				1.243
	50대				1.685
	60대				1.742
	70대 이상				2.393
성별	(남성)	1	7.8018	0.0052	
	여성				0.847
통신사	(SKT)	2	64.5263	<.0001	
	KT				0.836
	LG U+				0.496
연령*성별	(20대,남성)	5	22.4033	0.0004	
	30대				-0.621
	40대				-0.991
	50대				-0.730
	60대				-0.552
	70대이상				-1.688
권역*성별	(서울,남성)	6	14.1827	0.0277	
	경인				-0.819
	충청				-0.292
	전라				-1.266
	경북				-0.836
	경남				-0.452
강원제주	-0.543				

4.4 실험결과 요약

실험연구를 통해 확인한 전화면접조사와 ARS조사 간의 조사모드에 따른 차이와 유선/무선 RDD 전화번호와 가상번호 간의 표본추출틀에 따른 차이에 대해 간략히 요약하면 다음과 같다.

첫째, 조사항목의 응답분포에서는 표본추출틀 간 차이는 두드러지지 않지만 조사모드 간 차이는 뚜렷한 경향을 보였다. 특히 전화면접조사는 중도적 성향 및 모름/무응답의 응답이 상대적으로 많았던 반면, ARS조사는 양극단적 성향의 응답이 상대적으로 높게 나타났다. 예로,

- 정치적 관심도는 전화면접조사에서 약간 있거나 (별로 혹은 전혀) 없다는 경향이 높은 반면, ARS조사에서는 많이 있다는 경향이 매우 높게 나타난다.
- 정치적 성향에 있어서는 전화면접조사가 약간 진보적이거나 모르겠다/무응답이 많은 반면, ARS조사는 매우 보수와 매우 진보 (즉, 양극단) 측면에서 높은 비중을 보였다.
- 여론조사 효용성에 대해서는 전화면접조사는 대체로 믿을 만하거나 모르겠다/무응답의 비중이 높은 반면, ARS조사는 매우 믿을 만 하거나 (대체로 혹은 절대) 믿을 수 없다는 비중이 상대적으로 높게 나타났다.
- 대통령 국정수행의 평가에서는 전화면접조사는 대체로 잘하거나 대체로 잘못하고 있다는 중간적 응답이 많은 반면, ARS조사는 매우 잘한다거나 매우 잘못한다는 양극단적인 응답이 많았다.

둘째, 표본전화번호의 접촉과 분류에 있어서 전화면접조사는 ARS조사에 비해 보다 상세한 구분이 가능한 것으로 보인다. 예로, 전화면접조사는 거절과 중도이탈을 각각 구분할 수 있는 반면, ARS조사는 중도이탈만 확인할 수 있었다. 또한 전화면접조사는 부재중, 사업체번호, 팩스번호를 분류할 수 있었지만 ARS조사는 그러지 못하였다.

셋째, 전화면접조사와 ARS조사 모두 목표응답을 달성하기 위해 최초 표집하는

표본전화번호의 규모가 매우 큰데, 특히 전자에 비해 후자의 규모가 좀 더 크게 나타났다. 거절/중도이탈의 규모는 전화면접조사에 비해 ARS조사가 월등히 크고, 동일모드 내에서는 무선RDD/가상번호에 비해 유선RDD가 작았다. 하지만 결번은 유선 RDD가 무선 RDD에 비해 매우 많고, 통화중/부재중 등의 비접촉은 유선 RDD에 비해 무선 RDD가 높았다.

넷째, 응답성향모형 분석결과로 볼 때 가상번호와 유선 RDD에 대해서는 전화면접조사와 ARS조사 간에 응답성향은 다소의 차이도 존재한다. 전화면접조사는 권역, 연령, 통신사의 통계적 유의성이 있다면 ARS면접조사는 연령과 통신사는 물론 성별에 따른 권역과 연령의 상호작용도 유의한 것으로 나타났다.

실험연구는 제약적 상황에서 이루어져 조사결과의 해석은 다소 주의가 필요하다. 예를 들면, 전화면접조사 전문기관과 ARS 조사 전문기관이 보유하고 있는 유선 및 무선 RDD의 포괄범위와 품질은 물론 조사전략 및 운영방식도 서로 다를 수 있다. 다시 말해, 표본추출틀의 차이로 인해 발생할 수 있는 포함오차, 조사모드 및 전화시간과 재접촉 전략의 차이로 인해 발생하는 응답자 선택오차(selection bias), 동일 항목에 대한 조사모드에 따른 응답 간의 차이인 모드효과(mode effect) 등을 적절히 구분할 수 없었다.

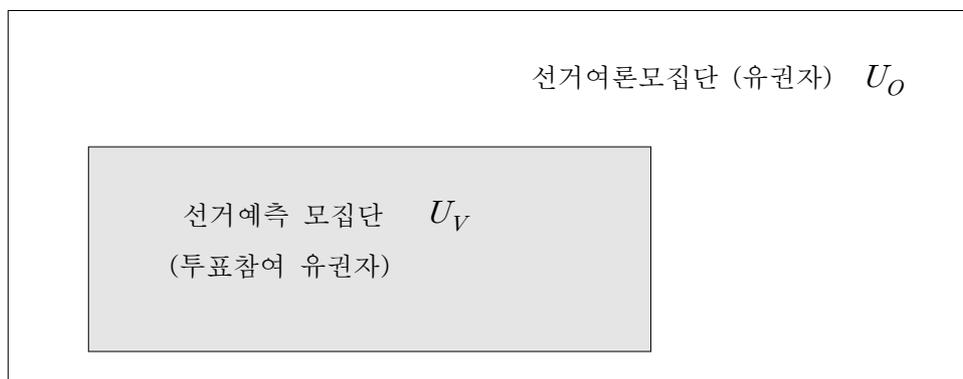
제 5 장 선거여론조사 개선 방안 논의

5.1 선거여론조사 방법의 현황과 문제점

1) 목표모집단: 여론파악과 선거예측

목표모집단(target population)이란 조사연구를 통해 추론하고자 하는 모든 기본 단위들의 집합으로 유한개의 단위로 구성되고 시간적 제약을 가지며 관측 혹은 접근이 가능하다. 선거여론조사는 여론을 파악하고자 하는 여론조사(public opinion poll)와 선거투표일에 참여한 유권자 선택의 결과예측(election forecast)의 두 가지 목적을 갖게 된다. 전자는 전체 유권자를 대상으로 하며 후자는 투표참여 유권자를 대상으로 하므로 후자는 전자의 일부분에 해당한다. <그림 5-1>은 조사목적에 따른 선거여론 조사의 목표모집단을 도식적으로 비교하고 있다. 더불어 선거여론조사의 조사시점과 선거투표일 간의 시간적 차이에 따라 여론은 물론 투표의향 또한 변할 수 있다. 따라서 특정시점에 진행되는 선거여론조사의 집계결과는 해석상 주의가 요구된다고 할 수 있다.

<그림 5-1> 선거여론조사 조사목적에 따른 목표모집단 비교



동일한 선거여론조사의 조사 자료를 통해 여론동향과악은 물론 선거예측을 함께 하기 위해서는 ‘투표의향’을 묻고 이에 따라 조사결과를 분석할 수도 있을 것이다. 하지만, ‘선거에 투표하실 의향이 있으십니까’와 같은 문항은 응답자로 하여금 소위 ‘사회적 바람직함(social desirability)’에 빠지게 하여 자신들을 좋게 보이도록 만드는 여과과정을 통해 대답할 수 있다(Babbie, 353쪽). 다시 말해, 투표의향이 과도하게 추정되는 측정오차(measurement error)가 발행할 수 있다. 또한, 조사시점과 선거시점 간 시간차이에 따른 유권자의 투표의향은 물론 ‘모르겠다/무응답’의 일종의 부동층의 변동은 선거예측 모집단에 대한 대표성과 관련하여 향후 연구가 필요할 것으로 판단된다.

2) 조사모집단: 다중표집틀 사용

조사모집단(survey population)이란 목표모집단에 대해 실제로 조사가 가능한 모집단을 뜻하며 일반적으로 표본추출틀(sampling frame) 혹은 이를 통해 접근이 가능한 모든 대상자들이 된다. 만약, 유권자 명부(voter list)가 있다면 이는 선거여론조사를 위한 최적의 조사모집단이 될 수 있을 것이다.

조사모집단 혹은 표본추출틀과 관련하여 포함오차와 정보오류의 두 가지 불안정성이 존재할 수 있다. 과소포함(undercoverage)과 과대포함(overcoverage)은 목표모집단의 일부 단위만이 혹은 목표모집단에 속하지 않는 단위가 표본추출틀에 과도하게 포함되는 경우에 발생한다. 예로, 선거여론조사에서 유선전화명부를 사용하면 전화번호가 등재되지 않았거나 전화가 없는 가구들은 조사에서 제외되는 과소포함이 발생하고, 팩스번호나 사업체번호들은 조사에 관심대상이 아니다.

동일 단위가 두 번 이상 표본추출틀에 기재되었다면 이는 중복기재(duplicate listing)에 해당한다. 예로, 한 사람이 다수의 전화회선을 보유한다면 그 만큼 표본에 포함될 가능성은 높아지게 된다. 또한 표본추출틀에 제공되는 보조정보가 부정확하다면 이로 인해 표본설계의 비효율성을 초래된다. 예로, 특정 번호가 20대 남성이 사용하는 것으로 기재되었지만 조사결과 50대 여성인 경우가 발생한다면 이는 추정량의 분산을 증가시키는 원인이 된다.

우리나라 선거여론조사에서는 주로 비용과 시간이 적게 드는 전화조사를 주로 활용하고 있다. 따라서 표본추출단위는 전화종류에 따라 유선전화, 무선전화 혹은 가상번호의 3가지로 나뉜다. 유선전화는 주로 개인보다는 가구가 공유하는 하므로 일종의 집락(cluster)으로 전화번호를 공유하는 전체 유권자를 포함한다. 유선전화는 2000년 이후 사생활 보호 등의 이유로 등재되지 않는 번호가 증가하게 되었고, 2008년 이후에는 전화번호부 인명편의 갱신 서비스가 중단되면서 기존 전화번호부는 사실상 표본추출틀의 역할을 하지 못하게 되었다. 이에 지역번호와 국번 이외의 마지막 네 자리를 무작위로 생성하여 이를 표본추출틀로 활용하는 유선번호 RDD(random digit dialing)가 2011년 재보궐선거에서 본격적으로 활용되기 시작하였다. <표 5-1>는 특정 조사회사들이 보유하고 있는 유선번호 RDD의 권역별 국번수를 비교하고 있다. A사와 B사는 각각 10,540개와 6,498개 국번수를 보유하고 있는데, C사의 4,336개에 비해 약 2.4배와 1.5배가 많음을 알 수 있다. 하지만 이들 유선번호 RDD 내에는 사업체번호, 팩스번호, 결번 등 가구전화가 아닌 비적격 회선이 상당수 포함되어 있으며, 더불어 유선 RDD의 유권자에 대한 포함률은 알려진 바 없다. 참고로 2018년 기준 유선전화 가입 회선수는 약 1,433만개이고 집전화 보유율은 38.4%(한국갤럽조사연구소, 2019)로 알려져 있다. 따라서 유선 RDD의 사용은 가구 유선번호의 낮은 등재율을 다소 극복할 수 있으나 낮은 재택률로 인해 주부, 고령층 등의 비경제활동인구가 과대하게 접근하게 되는 한계점을 갖는다.

<표 5-1> 유선번호 RDD 권역별 국번수

권역	국번수		
	A사	B사	C사
서울	2,885	1,551	1,097
경기, 인천	2,392	1,510	975
충청	1,087	750	450
전라	1,094	731	468
경북	874	610	450
경남	1,675	1,038	721
강원, 제주	533	308	175
합계	10,540	6,498	4,336

근이 가능한 개인은 그렇지 않은 개인에 비해 표본에 포함될 가능성은 높다. 확률추출방식에서는 이를 차별적인 추출확률을 적용하여 대표성을 표현한다. 물론 두 프레임 모두에서 선거여론조사에 적격하지 않은 번호가 존재할 수도 있다. 각 프레임에는 결번, 사업체번호 등 부적격한 번호들이 다수 존재하므로 전화조사를 통해 적격성을 일일이 확인해야만 하는 비효율성이 있게 된다. 더불어 무선 RDD는 사용 중인 거의 모든 무선전화 번호를 포함할 수 있는 가능성은 있지만, 번호 자체에 지역을 식별할 수 있는 정보가 없어 시도, 시군구, 선거구 단위조사에서는 사실상 활용이 불가하다.

2017년 2월 공직선거법 개정에 따라 공표, 보도용 선거여론조사에 한해 휴대전화 가상번호의 활용이 가능해졌다. 가상번호란 이동통신사 3개 회사 (SKT, KT, LG U+)가 보유한 실제 사용 중 무선전화번호들로 조사기관이 조사목적에 위해 여심위에 신청하면 가상의 번호형태로 제공받게 된다. 따라서 가상번호는 궁극적으로는 무선번호 프레임에 해당하지만, 거의 모든 번호가 적격 번호들로 이루어지고 이용자의 성별·연령·거주지역의 정보가 함께 제공받을 수 있어 조사수행에 있어 효율적이다. 단, 조사 시작일 최소 10일 전에 온라인을 통해 여심위에 신청하여야 하며, 1건당 1일 기준으로 16.83원(부가세 포함)을 지불해야 하며, 목표 표본의 최대 30배까지 통신사별로 나눠서 신청할 수 있다.

3) 응답성향: 무응답 편향

대부분의 표본조사에서는 무응답을 피할 수는 없으며 이로 인해 발생하는 오차, 즉 전체 표본개체 중 일부로부터만 자료를 얻게 되어 발생하는 오차를 무응답오차(nonresponse error)라고 한다. 무응답이 발생하면 이로 인해 (i) 응답수가 줄게 되어 표본추정량의 표본오차는 커지며, (ii) 응답과 무응답 개체 간에 존재하는 평균적 특성 차이는 잠재적인 편향(bias)인 비표본오차를 초래한다.

무응답이 발생하는 요인은 크게 두 가지로 구분할 수 있다. 첫째, 표본개체에 대해 접촉을 실패하였거나 접근이 불가능한 경우이다. 예로, 면접원이 유선전화를 걸었을 때 표본가구원이 집에 없거나, 조사기간의 제약으로 인해 더 이상 조사를 시도할 수 없을 때, 혹은 전화번호 정보가 부정확하거나 오래되어 쓸모없게 된 경우이다. 둘째, 표본개체에 대한 조사설득이 실패하는 경우이다. 예로, 조사대상자가 응답할 의사가

없거나, 전화조사에서 최초 수신자가 선택된 표본가구 내 가구원을 바꾸어 주지 않는 경우 등에 해당한다.

모수추정에 대한 무응답이 주는 영향력은 무응답오차 혹은 무응답편향의 형태로 평가될 수 있다(Valliant *et al.*, 2013). 예로, 개별개체가 조사에 협조하여 응답하는 것은 표본선택과 관계없이 이미 정해져 있어서 응답개체 s_R 와 무응답개체 s_{NR} 으로 나뉜다고 가정하는 고정응답모형(fixed response model, FRM) 하에서는 모평균 \bar{Y} 의 Hajek-평균추정량 \hat{y}_π 의 편향이 다음과 같이 정의될 수 있다.

$$bias(\hat{y}_\pi) = \frac{N_{NR}}{N}(\bar{Y}_R - \bar{Y}_{NR}) \quad (5-1)$$

여기서 Hajek-평균추정량은 $\hat{y}_\pi = \sum_{S_R} d_{oi} y_i / \sum_{S_R} d_{oi}$ 이고, d_{oi} 는 표본추출률의 역수인 설계가중치(design weight)이고, $N = N_R + N_{NR}$ 은 모집단 크기, N_R 와 N_{NR} 은 각각 모집단 내 전체 응답과 무응답 개체수, \bar{Y}_R 와 \bar{Y}_{NR} 은 각각 응답 및 무응답 개체의 모평균을 나타낸다. 즉, 무응답이 결정적 현상이라면 평균추정량의 무응답편향은 (i) 응답자와 무응답 간의 평균값 간의 차이가 크거나, (ii) 무응답률이 클수록 커질 수 있다는 것이다. 반면, 조사요청 시 개체별 협조여부는 확률적으로 결정된다고 가정 하는 확률 응답모형(random response model, RRM) 하에서는 Hajek-평균추정량의 편향은 다음과 같이 유도된다.

$$bias_{RRM}(\hat{y}_\pi) = \frac{1}{\bar{\phi} N} \sum_{i \in S} (y_i - \bar{Y}_U)(\phi_i - \bar{\phi}) \quad (5-2)$$

여기서 $\bar{\phi} = N^{-1} \sum_U \phi_i$ 는 식(4-1)에서 정의하는 응답확률의 모평균을 나타낸다. 무응답편향은 조사변수 y 와 응답확률 ϕ 의 공분산과 응답확률의 모평균 역수에 비례한다. 따라서 y 와 ϕ 간에 상관관계가 없다면 무응답편차도 발생하지 않는다.

<표 5-2>와 <표 5-3>은 4장의 실험연구에서 각각 가상번호에 대해 전화면접 조사와 ARS조사를 적용하여 확인한 응답 및 무응답수, 제공정보 일치율, 응답산출률 및 추출배수를 성별, 연령별 및 권역별로 정리하고 있다. 여기서 응답산출률이란 분류 특성 내 전체 가상번호 대비 응답번호의 상대크기, 즉 비율을 나타낸다. <표 4-10>에서 확인할 수 있듯이 가상번호에서 미사용을 제외한 비적격 번호의 규모가 매우 작

으므로 응답률과 거의 같다고 볼 수 있다. 따라서 식 (5-1)의 고정응답모형으로 무응답 편향을 평가한다면 전화면접조사에서 남성의 경우에 응답(산출)률이 7.5%이므로 응답자 특성과 무응답 특성이 조금만 차이가 나더라도 편향은 매우 클 수 있음을 나타낸다. 예로, 응답자와 무응답자의 대통령 국정지지율의 차이가 10%일 때 무응답(산출)률 92.5%(=100%-7.5%)은 응답자만으로 보고되는 표본추정량의 무응답 편향이 약 9.25%로 매우 클 수 있음을 알 수 있다. 남성의 ARS 조사 응답(산출)률은 4.8%로 더욱 낮기 때문에 표본추정량의 무응답 편향은 약 9.52%로 더욱 커짐을 알 수 있다. 반면, 확률응답모형으로 무응답 편향을 평가한다면 응답확률과 조사특성 간의 상관관계가 높을수록 커짐을 알 수 있다.

또한 가상번호에서 제공하는 전화번호의 특성의 정확도에 대해서 응답자에 한해서 확인할 수 있다. <표 5-2>와 <표 5-3>을 살펴보면 전화면접조사에 비해 ARS 조사에서 응답자들에 대한 제공정보(즉, 성, 연령, 지역)의 일치율은 다소 낮게 나타난다. 물론 이러한 정보 불일치는 통신사에서 제공한 정보가 잘못된 경우도 있겠으나 응답자가 고의 혹은 단순한 실수에 의해 잘못된 정보를 제공하였을 수도 있다.

한편, 가상번호 신청배수는 ARS조사의 경우에는 19~29세의 연령대, 여성, 전라권에 대해서는 30배가 넘는 것으로 나타난다. 예로, 19~29세 유권자 한명의 응답을 얻기 위해 역 42개의 가상번호에 대해 조사접촉의 시도를 해야 함을 의미한다.

<표 5-2> 전화면접조사 가상번호 제공정보 일치율 및 추출배수

제공정보		응답	일치	불일치	무응답	전체	일치율	응답 산출률	추출 배수
성별	남	211	199	12	2,596	2,807	94.3%	7.5%	13.3
	여	197	182	15	2,981	3,178	92.4%	6.2%	16.1
	합계	408	381	27	5,577	5,985	93.4%	6.8%	14.7
연령대	19~29세	75	72	3	1,147	1,222	96.0%	6.1%	16.3
	30~39세	72	63	9	1,141	1,213	87.5%	5.9%	16.8
	40~49세	66	59	7	1,052	1,118	89.4%	5.9%	16.9
	50~59세	78	65	13	976	1,054	83.3%	7.4%	13.5
	60~69세	61	49	12	665	726	80.3%	8.4%	11.9
	70세 이상	56	40	16	596	652	71.4%	8.6%	11.6
	합계	408	348	60	5,577	5,985	85.3%	6.8%	14.7
권역	서울	76	65	11	1,044	1,120	85.5%	6.8%	14.7
	경기,인천	130	122	8	1,638	1,768	93.8%	7.4%	13.6
	충청	39	38	1	691	730	97.4%	5.3%	18.7
	전라	33	31	2	501	534	93.9%	6.2%	16.2
	경북	38	36	2	526	564	94.7%	6.7%	14.8
	경남	70	66	4	857	927	94.3%	7.6%	13.2
	강원,제주	22	18	4	320	342	81.8%	6.4%	15.5
	합계	408	376	32	5,577	5,985	92.2%	6.8%	14.7

<표 5-3> ARS조사 가상번호 제공정보 일치율 및 추출배수

제공정보		응답	일치	불일치	무응답	전체	일치율	응답 산출률	추출 배수
성별	남	343	326	17	2,596	7,214	95.0%	4.8%	21.0
	여	217	188	205	2,981	6,967	86.6%	3.1%	32.1
	합계	560	514	222	5,577	14,181	91.8%	3.9%	25.3
연령대	19~29세	98	76	22	3,956	4,054	77.6%	2.4%	41.4
	30~39세	120	73	47	3,492	3,612	60.8%	3.3%	30.1
	40~49세	101	78	23	2,665	2,766	77.2%	3.7%	27.4
	50~59세	105	79	26	1,761	1,866	75.2%	5.6%	17.8
	60~69세	79	54	25	1,079	1,158	68.4%	6.8%	14.7
	70세 이상	77	51	26	648	725	66.2%	10.6%	9.4
	합계	580	411	169	13,601	14,181	70.9%	4.1%	24.5
권역	서울	109	97	12	2,344	2,453	89.0%	4.4%	22.5
	경기,인천	161	144	17	4,465	4,626	89.4%	3.5%	28.7
	충청	64	58	6	1,473	1,537	90.6%	4.2%	24.0
	전라	40	35	5	1,331	1,371	87.5%	2.9%	34.3
	경북	55	43	12	1,344	1,399	78.2%	3.9%	25.4
	경남	78	63	15	2,116	2,194	80.8%	3.6%	28.1
	강원,제주	36	29	7	565	601	80.6%	6.0%	16.7
	합계	543	469	74	13,638	14,181	86.4%	3.8%	26.1

4) 파라데이터: 자료공개, 투명성, 조사오차 평가

조사품질을 결정하는 중요한 하나의 요인은 투명성(transparency)인데 이는 정보 이용자가 조사 실시방법 및 결과 도출 방법에 대해 얼마나 투명하게 알 수 있도록 정보가 주어지는지에 따라 결정된다. <표 5-4>는 미국 CNN이 조사결과를 인용할 때 점검하는 16가지 질문⁴⁾을 나열하고 있다. <표 5-5>는 여심위에 지정하는 등록 및 공표자료에 공개해야 되는 파라데이터의 기준⁵⁾을 나열하고 있다. CNN의 질문과 비교할 때 우리나라 선거여론조사에서 제공하는 정보도 상당히 많다는 것을 확인할 수 있다. 하지만 상세하게 살펴보면 몇 가지 측면에서 우리나라 선거여론조사에서 제공하는 정보의 개선점이 발견된다.

첫째, 재접촉 시도에 대한 정보 제공 및 공개 정도이다. <표 4-11>과 <표 4-14>의 비교에서 추측할 수 있듯이 일반적으로 수행되는 선거여론조사에서 충분한 재접촉이 수행되는지에 대해 논의 사례는 그리 많지 않은 것으로 판단된다.

둘째, 조사오차의 공표가 주로 응답률 100%를 가정하는 표본오차로 이루어지며 설계효과나 총조사오차의 관점에서의 비표본오차에 대한 연구는 부족한 것으로 판단된다. 3장에서 살펴본 바대로 우리나라 선거여론조사는 오로지 응답수에 의해서만 표본오차를 보고하고 있다.

셋째, 가중치 산출이 현재 할당특성에 따른 모집단 비율과 표본응답자 비율을 맞추는 작업으로만 이루어지고 있다. 다시 말해, 표본전화번호의 추출확률 및 무응답 조정 등과 같은 관점에서의 단계별 상세한 절차는 생략되고 있다. 또한 CNN의 점검 질문에 명시된 교육수준과 같은 추가적 주요변수에 대한 무응답 편향 가능성에 대한 논의는 부족한 것으로 판단된다.

4) <https://verstaresearch.com/blog/how-polls-pass-cnns-quality-review/>

5) <https://www.nesdc.go.kr/portal/content/view.do?menuNo=300007>

<표 5-4> CNN 조사연구의 품질 투명성 평가 질문

1. 어느 조사기관이 여론조사를 실시하였는가?
2. 여론조사는 전화면접, ARS, 온라인, 자기기입식 설문지 혹은 다른 방법들 중 어느 것이 적용되었는가?
3. 조사비용은 누가 지불했고 왜 그리 되었는지?
4. 여론조사에 얼마나 많은 사람들이 인터뷰했는가?
5. 응답자는 어떠한 언어로 조사되었는가?
6. 여론조사에 포함된 모든 질문문항에 대한 지침과 프로그래밍에 관련한 모든 내용을 공개할 것.
7. 언제 조사가 수행되었는가?
8. 표본추출틀은 무엇이고 응답자는 어떻게 선택되었는지에 대해 가능한 상세히 기술할 것. 만약 웹패널(들)이 사용되었다면, 어떻게 패널들이 모집되었는지 기술할 것. 여론조사가 온라인으로 진행되었고 라우터(router)⁶⁾를 통한 참여자 섭외가 되었다면 그 비율은 얼마나 되는가?
9. 만약 표집할당 혹은 응답자할당이 적용되었다면 어느 단계에서 적용되었고, 할당변수 혹은 목표가 사용되었으며, 목표할당(target)에 대한 추정에 쓰인 자료는 무엇인가?
10. 목표모집단은 무엇이고 선택된 표본추출틀이 목표모집단을 대표한다고 어떻게 신뢰할 수 있는가?
11. 전화면접조사가 사용되었다면 얼마나 많은 부분이 휴대전화로 진행되었는가? 온라인조사이었다면 이동통신기기(mobile browser)가 허용되었는가? 그렇다면 어느 정도나 이동통신기기로 응답하였는가?
12. 전화면접조사가 사용되었다면, (최종적 제외 전에) 몇 번의 재접촉이 시도되었는지?
13. 조사진행이 면접방식(live interview)이 아니었다면, 조사대상이 실제 사람인지 또한 조사대상이 조사에 주목하고 있는지에 대해 확신할 수 있는가?
14. 조사오차(survey error)는 어떻게 정의하고 추정하는가? 그리고 제시한 조사오차가 적절하다고 판단하는 근거는 무엇인가? 만약 조사에 대한 (표본)오차한계(margin of sampling error)가 보고하고 있다면, 설계효과(design effect)⁷⁾에 대한 조정이 포함되고 있는가?
15. 조사자료 분석에 가중치가 사용되었다면, 해당 가중치 산출에 사용된 변수들과 가중치 모수 및 산출과 관련된 자료출처를 밝힐 것. 만약 조사결과산출에서 응답자 교육수준이 조정되지 않았다면, 그 이유를 설명하고 교육수준에 대한 비가중 분포를 제시하라.

6) 라우터에 관한 내용은 다음의 사이트를 참고할 수 있다. <https://www.elitesurveysites.com/survey-routers/>

7) 설계효과란 조사연구에 실제로 적용된 표본설계 및 조사결과에 대한 추정량이 갖는 분산값을 단순표집 하에서 갖는 (단순한 형태의 비가중) 추정량이 갖는 분산값으로 나눈 비(ratio)로 정의된다.

<표 5-5> 중앙선거여론조사심의위원회 등록 및 공표자료

등록사항	함께 공표·보도하여야 할 사항
<ol style="list-style-type: none"> 1. 선거여론조사의 명칭 2. 조사의뢰자 3. 선거여론조사기관 4. 조사지역 5. 조사일시 6. 조사대상 7. 조사방법 8. 표본의 크기 9. 피조사자 선정방법 (휴대전화 가상번호 사용 시 그 사실 및 사용비율 포함) 10. 피조사자 접촉 현황 (비적격 사례수는 결번과 그 외의 비적격 사례로 구분하여 등록한다.) 11. 접촉률 12. 응답률 13. 가중값 산출 및 적용 방법 14. 표본오차 15. 전체 질문지 16. 결과분석 17. 최초 공표·보도 예정일시 	<p>최초 공표·보도</p> <ol style="list-style-type: none"> 1. 조사의뢰자 2. 선거여론조사기관 3. 조사지역 4. 조사일시 5. 조사대상 6. 조사방법 7. 표본의 크기 8. 피조사자 선정방법(휴대전화 가상번호 사용 시 그 사실 및 사용비율 포함) 9. 응답률 10. 가중값 산출 및 적용 방법(DB, 패널 등 특정 표본추출틀 사용 시 전체 규모와 구축방법 포함) 11. 표본오차 12. 질문내용 <p>인용 공표·보도</p> <ol style="list-style-type: none"> 1. 조사의뢰자 2. 선거여론조사기관 3. 조사일시 4. 그 밖의 사항은 중앙선거여론조사심의위원회 홈페이지 참조 <p>다수의 선거여론조사 분석결과 공표·보도</p> <ol style="list-style-type: none"> 1. 분석의뢰자 2. 분석기관·단체 3. 분석대상(기간,건수,출처) 4. 분석방법 5. 각 여론조사 내용은 중앙선거여론조사심의위원회 홈페이지 참조

5.2 선거여론조사 방법의 개선방법 모색

1) 확률추출로의 변환

비확률표집방법인 할당추출을 기초로 진행하는 현재의 선거여론조사가 확률표본 방식으로 전환하기 위해서는 표본이론에 따른 확률추출방식을 따라야 할 것이다. 이를 위해서는 표본전화번호의 표본설계가 확률추출방식이어야 하며 더불어 표본추출률이 설계가중치로 계산이 가능하여야 할 것이다. 무선 RDD를 사용하는 경우는 딱히 번호별 특성을 알 수 없어 단순확률추출에 해당하며 이러한 측면이 적용되어야 한다.

반면, 유선 RDD나 가상번호를 사용하는 경우에는 전자는 지역구분과 후자는 통신사는 물론 성·지역·연령으로 정해지는 표본추출률이 계산되어야 한다. 현재 통신사들이 영업비밀이라는 이유로 관련 정보를 제공하고 있지 않아 모집단 구성정보, 즉, 각 통신사별 가입자들에 대한 (성·지역·연령) 특성별 총 번호수의 정보를 확인할 수 없다. 이러한 정보가 제공된다면 확률추출법의 하나인 층화확률추출의 적용이 가능해진다. 4장의 실험연구에서 확인한 바와 같이 표본추출이 통신사별로 이루어지고 또한 응답성향도 통신사별로 유의미한 차이가 발생하므로 ‘통신사’는 표본층 (혹은 적어도 무응답 조정층)의 기준으로 고려되어야 할 것이다.

[대안 1] 가상번호만을 사용한다면 현행방식을 수정하여 확률추출을 실시한다면 다음과 같이 표본설계 및 가중치 산출을 구성을 변경할 수 있다.

구성 1) 층화확률추출

통신사 및 할당특성은 모집단 층화에 해당하며 중앙선거여론조사심의위원회를 통해 신청하는 번호수 n_h 는 해당층의 총 번호수 N_h 와 비교하여 다음과 같이 설계가중치를 계산된다. (단, 통신사에서 해당 층의 정보를 제공할 수 있어야 한다.)

$$w_{hi} = \frac{N_h}{n_h}$$

구성 2) 무응답조정

통신사에서 제공하는 가상번호에 대한 보조정보 x_{hi} 를 이용하여 다음과 같은 무응답조정을 실시한다.

$$w_{hi}^{NR} = \frac{w_{hi}^D}{\phi(x_{hi})}$$

여기서 응답성향 $\phi(x_{hi})$ 은 식 (4-1)의 모형식으로 로지스틱회귀모형 혹은 프로빗모형 등을 고려하여 적합할 수 있다.

구성 3) 사후층화조정

모집단 총합 정보를 이용하여 사후층(post-stratum)의 가중치 합이 모 총합과 일치하도록 조정한다.

$$w_{hi}^{PS} = w_{hi}^{NR} \frac{N_g}{\hat{N}_g} \quad ((hi) \in U_g (g=1, \dots, G)) \quad (5-3)$$

여기서 U_g 와 N_g 는 사후층과 사후층 총수, $\hat{N}_g = \sum_{(hi) \in s \cap U_g} w_{hi}^{NR}$ 는 사후층에 포함된 표본개체들의 무응답조정 가중치 총합을 각각 나타낸다.

표본가중치가 커서 자료분석에 과도한 영향을 줄 수 있다면 이는 이상치(outlier)이고 가중치절사(weight trimming)을 통해 그 영향력을 적절히 줄일 수 있다.

[대안 2] 유무선 병합 혹은 유선과 가상번호를 병합하여 조사한다면, 유선번호의 집락특성과 동일추출틀 내 중복(duplication) 혹은 다중추출틀(multiframe) 사용에 따른 중복포함의 가능성을 고려한 개인별 추출확률은 다음의 구성을 고려하여 계산되어야 한다.

구성 1) 표본추출확률

가구단위로 사용되는 유선전화번호의 경우, 가구 내 복수 유권자가 접촉될 가능성을 반영하여 다음과 같이 설계가중치를 산출한다.

$$w_{hi}^d = M_{hi} \left(\frac{N_h}{n_h} \right)$$

여기서 구성요소 N_h/n_h 는 해당 유선번호가 표본에 포함될 포함확률의 역수이고 M_{hi} 는 표본가구 내 총 유권자 수를 나타낸다. 유권자수가 과도하게 많은 가구인 경우에는 해당 가구 내 유권자 선택이 주는 과도한 가중치 영향력을 줄여주기 위해 최대 2가 넘지 않도록 절사하여 다음과 같이 개인 표본설계가중치를 산출할 수 있다.

$$w_{hi}^d = \min(2, M_{hi}) \times \left(\frac{N_h}{n_h} \right)$$

무선전화와는 달리 유선전화는 지역특성을 구분할 수 있어 이를 표본층으로 간주할 수도 있다. 가구 내 조사대상 선택은 단순확률추출법 혹은 유사확률(예, 생일법)에 따라 선택할 수 있다.

구성 2) 중복제거 혹은 중복포함확률

표본추출틀 내 특정 가구 혹은 개인이 중복되어 포함될 수 있다. 동일 표본추출틀 내 중복개체를 먼저 제거한 후 표본추출을 실시하는 것이 바람직하지만 그렇지 못하다면 전화조사를 통해 중복번호의 가능성을 조사하여 확인하여 표본포함확률을 조정할 수 있다. 예로, 표본추출틀에 총 m_i 개의 동일대상 전화번호가 존재한다고 가정하자. 표본추출이 서로 독립적이고 표본개체 i 에 대한 j 번째 나열 개체가 갖는 포함확률(inclusion probability)이 π_{ij} 이라면, 표본개체 i 의 추출확률은 다음과 같이 계산할 수 있다.

$$\pi_i = 1 - \prod_{j=1}^{m_i} (1 - \pi_{ij}) \quad (5-4)$$

구성 3) 다중추출틀 표본설계 반영

논의를 단순화하기 위해 두 개의 표본추출틀 U_A 와 U_B 만을 사용하는 표본대상을 선택한 <그림 5-2>의 상황을 가정하자. 개별 표본추출틀에만 존재하는 대상은 각각 π_{iA} 와 π_{iB} 의 포함확률만 존재한다. 하지만 표본추출틀 모두에 속하는 대상은 앞서 개별 포함확률이 종합적으로 고려(예로, 식 (5-4))되어 정의된다. 이러한 측면은 [대안 1]에 기술된 표본가중치

산출이 각각의 표본추출에 적용된 후 절충적으로 반영될 수 있다. 아주 간단히 기술한다면 다음과 같은 방식으로 표본가중치를 산출할 수 있다.

$$w_i = \begin{cases} w_{iA} & i \in s_a (\subset U_a) \\ f(w_{iA}, w_{iB}) & i \in s_{ab} (\subset U_{ab}) \\ w_{iB} & i \in s_b (\subset U_b) \end{cases} \quad (5-5)$$

여기서 표본추출틀 A는 해당 추출틀에만 있는 번호와 표본추출틀 B에도 함께 존재하는 번호로 $U_A = U_a \cup U_{ab}$ 와 같이 나뉘며, 표본추출틀 B도 유사하게 $U_B = U_b \cup U_{ab}$ 로 나뉜다고 가정하다. 또한 표본도 마찬가지로 $s_A = s_a \cup s_{ab}$ 와 $s_B = s_b \cup s_{ab}$ 으로 나뉠 수 있다. 예로, 유무선 병합인 경우에는 s_a 는 유선번호로만 접근이 가능한 유권자, s_{ab} 는 유선과 무선번호로 모두 접근이 가능한 유권자, s_b 는 무선번호로만 접근이 가능한 유권자를 나타낸다. 식 (5-5)에 의해 산출되는 다중표본틀을 사용한 표본설계에 대한 표본가중치 산출은 Skinner(1991)를 참고할 수 있다.

현재 우리나라 선거여론조사자료(3장 참고)에서 전체 표본수(응답, 무응답, 부적격을 모두 포함)와는 상관없이 응답자 중 할당특성비율이 주민등록상 유권자의 할당 특성비와 같게 놓는 비례식, 즉 $n_q^R/n^R = N_q/N$ 이 되도록 하되, 만약 이러한 조건을 만족시키지 못한다면 추가적으로 다음의 조정을 실시한다.

$$a_q \frac{n_q^R}{n^R} = \frac{N_q}{N} \quad (5-6)$$

여기서 a_q ($0.7 \leq a_q \leq 1.5$)는 할당특성 q (예, 성, 연령, 지역)에 대한 조정계수⁸⁾이다. N 과 N_q 는 각각 전체 유권자와 특정 할당특성을 갖는 유권자, n^R 과 n_q^R 은 각각 전체 응답자수와 특정 할당특성을 갖는 응답자수를 나타낸다. 할당특성 q 를 갖는 응답자 i 에 대한 표본가중치는 다음과 같이 표기될 수 있다.

$$w_i^{Quota} = \frac{N_q}{n_q^R} = \frac{a_q N}{n^R} \quad (i \in U_q) \quad (5-7)$$

할당균형을 반영한 표본가중치 (5-7)은 표본추출의 단계 및 집락적 구조, 무응답조정, 중복포함확률, 다중추출틀 사용에 따른 절충 등의 다양한 요소들을 적절히 반영되

8) 가중값 배율조정에 관한 규정은 다음의 중앙선거여론조사심의위원회 홈페이지 내 관련법규를 참고할 수 있다.
<https://www.nesdc.go.kr/portal/content/view.do?menuNo=300010>

지 못하고, 식 (5-3)과 같이 할당특성에 따른 사후층화와 유사한 조정만이 고려되었음을 알 수 있다.

2) 조사비용 현실화

조사품질은 지속적인 노력, 시간, 예산 투입이 필요하며 우선순위를 두고 열정을 쏟아야 하며, 품질향상 노력에 대한 평가와 보상이 적절한 형태로 꾸준히 이루어져야 바람직하다. 품질향상을 저해하는 장애물 (예, 부적합 예산 집행, 여론조사에 대한 과도한 제도적 규제 등)을 헐고 시의적절한 피드백을 받아 추가적인 지원, 훈련, 방침 이행이 필요하다.

한국통계학회는 한국조사협회의 28개 회원사를 대상으로 2019년 4월 기준 조사 방법에 따라 표본 당 (i) 현재 평균 단가와 (ii) 적정 평균 단가를 조사하였다. 전국 성인 남녀 1,000명에 대해 응답시간 10분 내외의 조사를 기준으로 하는 CATI/RDD 유선전화와 무선전화의 현재 표본당 평균 단가는 각각 15,792원과 16,625원으로 적정 평균 단가 19,750원과 20,667원과 3,958원과 4,042원의 차이를 보였다. 최대 적정 평균 단가는 35,000원과 38,000원으로 현재 평균 단가는 매우 낮게 책정되어 있음을 알 수 있다. 향후 선거여론조사에 대한 보다 상세한 비용구조 및 비용요소를 파악한다면 적절한 수준의 품질을 갖춘 조사수행이 가능할 수 있을 것이다.

3) 비표본오차 프로필

비표본오차는 모든 조사단계에서 다양한 이유로 발생할 수 있다. 더불어 발생 가능한 모든 종류의 비표본오차를 해결할 수 있는 완벽하거나 특별한 이론이 존재하지 않는다. 다만 현재로서 고려할 수 있는 최선의 방법은 각각의 조사에 대해 “오차프로필(error profile)”, 즉 비표본오차의 주요 발생 원인들에 대해 기술하는 것부터 시작하는 것이다.

<표 5-6>은 총조사오차와 관련한 쟁점과 관련하여 Whitely (2014)이 고려한

오차 프로필의 예시를 나타내고 있다. 개별 조사별로 조사주기 단계별로 발생 가능한 오차들에 대해 점검 논의와 관련된 쟁점을 나열하고 이와 관련하여 조사통계 품질에 대한 잠재적 영향력을 평가함으로써 개별 오차에 대한 제어가 가능케 된다.

4) 향후 선거여론조사의 지속적 품질향상 노력 및 추가연구에 관한 제언

Marker and Morganstein(2004)는 국가통계 기준의 조사통계에 대하여 지속적 품질향상이 가능할 수 있는 필수요건을 <표 5-7>과 같이 나열하였다. 이는 우리나라 선거여론조사의 개선을 위한 노력의 방향을 잡는데 많은 시사점을 준다고 할 수 있다. 먼저, 선거여론조사와 관련된 각 주체(예, 조사기관, 정치권, 언론, 유권자, 중앙선거여론조사심의위원회, 통신사, 학계 등)별로 조사연구에 대한 이해도를 높이고 관련 전문가를 충분히 육성하고 확보하여야 할 것이다. 선거여론조사의 신뢰도에 대한 꾸준한 의문이 제기됨에도 불구하고 저비용 구조, 전문가 부족, 인프라 및 조사환경의 열악함 등에 대해 평가하고 이해하며 개선하려고 노력하기 보다는 조사결과의 단순 사용, 오용 및 남용 등을 통해 신뢰도를 더욱 떨어뜨린다고도 할 수 있다. 보다 신뢰할 수 있는 선거여론조사가 되기 위해서는 재정(조사단가)과 인력(전문가)에 대한 충분한 투자는 물론이고 선거여론조사에 대한 주체별 이해증진 및 품질에 따른 조사자료의 취사선택(통계교육), 인프라 개선(양질의 표본추출틀 구축, 응답률 제고), 주체별 신뢰구축 등이 지속적으로 관심을 갖고 행해져야 할 것이다.

선거여론조사의 객관성·신뢰성 제고를 위해 향후 고려할 수 있는 추가연구로는 (i) 가상번호, 유선RDD, 무선RDD 등 표본추출틀의 포함오차 평가, (ii) ‘투표의향’, ‘지지정당’ 등 조사문항에 대한 측정오차 평가, (iii) 조사모드에 따른 응답차이(모드효과), (iv) 적절한 조사비용, (v) 무응답 연구 등을 고려할 수 있을 것이다.

<표 5-6> 조사통계 총조사오차 및 잠재적 영향 파악을 위한 체크리스트 (예시)

	총조사오차 점검 논의	관련 쟁점	자료품질에 대한 잠재적 영향
대표성 관련 오차			
포함오차	조사범위 모집단이 어떻게 정의되었는가? 표본추출들의 구성이 조사적격 모집단과 일치하는가? 표본추출들에 부적격 혹은 중복된 개체들은 없는가?	조사범위 모집단이 부정확하거나 불량하게 정의됨 선택된 표본추출들이 우리나라 기업체를 적절히 대표하지 못함 부적격 개체들이 표본으로 추출됨	높음
표본오차	표본크기가 적절한가? 오차한계가 기대치보다 큰가?	표본크기가 부적절함 조사자료가 분석/보고 목적을 이루기 위해 불충분함	높음
무응답오차	단위무응답은 어떻게 정의되는가? 모집단 내 특정 하위집단의 무응답이 존재하는지? 항목무응답은 어느 정도인지?	무응답률이 높아 무응답 편향이 발생할 수 있음 특정 하위집단이 지나치게 많이 무응답으로 탈락함 특정 항목들에 대한 높은 항목무응답으로 편이 발생 가능	높음
조정오차	표본가중치 산출방식이 적절한가?	가중치 산출방식이 다소 부정확함	낮음
측정 관련 오차			
타당성	조사도구가 타당한지, 즉 신뢰할 수 있는지?	조사도구가 희망하는 개념을 측정하지 못하거나 일관적이지 못함	낮음
측정오차	설문지 설계가 잘되었는가? 조사원의 무의식적/의식적으로 부정확한 정보를 제공가능?	잘못된 설계로 부정확하거나 불안정한 응답을 받게될 가능성 있음 조사원이 응답자로 하여금 응답값을 변경하거나 조작할 수 있음	중간
처리오차	분석단위가 적절히 정의되었는지? 자료정제(data clean)이 어떻게 이루어지는지? 자료코드가 적절히 정의되었는지?	분석단위가 부적절히 정의됨 응답자료 타당성이 부적절히 확인 코딩 오류가 존재하거나 개방형 문항의 응답사항을 잘못 입력함	중간
추정오차	조사자료가 바르게 분석되었나?	옳지 않은 분석기법이 사용됨 부정확한 추정 발생	낮음

<표 5-7> 조사통계의 지속적 품질향상을 위한 필수요건

- 첫째, 고위관리층부터 지속적인 품질향상(continuous quality improvement, CQI)에 대하여 업무상 우선순위에 두고 열정을 쏟아야 할 것이다. 이러한 관리자의 리더쉽이 구현되려면 품질향상 노력에 대한 평가 및 보상이 적절한 형태로 꾸준히 이루어지는 것이 바람직하다.
- 둘째, 조직 구조적 측면에서 조사통계의 정확성을 향상시키기 위해서는 측정 및 무응답 오차와 관련된 이론적 훈련을 받은 전문가를 확보하여야 한다. 관련된 연구는 물론 통계업무상의 측정 및 무응답 오류를 이해하고 극복할 수 있는 능력을 배양하여야 하며, 다양한 형태와 특성의 조사오류에 대한 기관 안팎의 구성원들의 이해를 제고하여야 한다.
- 셋째, 지속적 품질개선에 대한 기관인식 혹은 문화는 반드시 자립적이고 자급적이어야 한다. 지도팀을 구성하여 지침 및 중간단계 교정 등을 제공하게 함으로 지속적 품질향상의 초기단계의 성공적 정착이 될 수 있도록 한다. 이를 위해 지도팀의 구성원은 최고관리층을 대표할 수 있는 자, 중간관리자, 업무수행자 등으로 구성하여 정기적 회의를 통해 품질향상을 저해하는 장애물을 헐고, 또한 시의적절한 피드백을 받아 추가적인 지원, 훈련, 방침 이행이 필요한 지를 결정할 수 있어야 한다.
- 넷째, 지속적 품질향상은 조직 내 탁월한 의사소통이 전제되어야 한다. 현행최상 방법론(current best methods CBM)과 같은 품질향상을 위한 최선의 방법론을 숙지하고 적용할 수 있도록 기관 간 업무 및 경험을 공유할 수 있어야 한다.
- 다섯째, 품질향상팀에 초급자로부터 전문가까지를 아우르고 조사통계의 세부업무별 수요자와 공급자를 포함할 수 있도록 한다. 이를 통해 조사절차에 대한 이해로부터 시작하여 업무상 투입물과 산출물은 물론 세부절차에 대한 지식을 갖는 사람들이 관여할 수 있도록 한다.
-

참고문헌

- 한국갤럽조사연구소 (2019). “전화조사 방법 및 이슈”.
- Babbie, E. (2014). *The Practice of Social Research* (13th ed.). Cengage. [고성호 외 역(2013) 『사회조사방법론』. 세계이자리닝코리아.
- Biemer, P. (2010). “Total Survey Error: Design, Implementation and Evaluation,” *Public Opinion Quarterly*, 74, pp. 817–848.
- Cochran, W. (1977). *Sampling Techniques*. John Wiley & Sons, Inc. New York.
- Groves, R.M. and Lyberg, L. (2010). Total Survey Error: Past, Present, And Future. *Public Opinion Quarterly* 74, 849–879.
- Lohr, S.L. (1999) *Sampling: Design and Analysis*. Duxbury Press, Pacific Grove, CA.
- Marker, D. A. and Morganstein, D. R. (2004). “Keys to Successful Implementation of Continuous Quality Improvement in a Statistifcal Agency,” *Journal of Official Statistics*, 20, pp. 125–136.
- Park, M. (2006). Alternative designs for Regression Estimation. *Journal of Official Statistics*. 22, 541–563.
- Rosenbaum, P. and Rubin, D. B. (1988). “The central role of the propensity score in observational studies for causal effects,” *Biometrika*, 70, pp. 41–45.
- Scali, J., Testa, V., Kahr, M., Strudler, M. (2005). “Measuring Nonsampling Error in the Statistics of Income Individual Tax Return Study,” *Proceedings of the Section on Survey Research Methods, American Statistical Association*.
- Skinner, C. J. (1991). “On the Efficiency of Raking Estimation for Multiple Frame Surveys,” *Journal of the American Statistical Association*, 86, pp. 779–784.
- Valliant, R., Dever, J. A., Kreuter, F. (2013). *Practical Tools for Designing and weighting survey samples*. New York: Springer.
- Valliant, R., Dorfman, A. H. and Royall, R. M. (2000). *Finite Population Sampling and Inference: A Prediction Approach*. Wiley.

Whitely, S. (2014). "Total Survey Error and Institutional Research: A Case Study of the University Experience Survey," presented at the Australasian Association for Institutional Research Forum, Melbourne, Australia.

부록 1. ARS 조사용 설문지

「주요 정치현안 관련 여론조사」	표본수	N= 1,000
--------------------------	-----	----------

조사지역	전국	조사방법	ARS전화조사□유선전화RDD□무선 가상번호
------	----	------	-------------------------

00. [인사말] 안녕하십니까? 여론조사 전문기관인 모노커뮤니케이션즈/모노리서치입니다. 저희 기관에서는 최근의 주요 사회현안에 대한 여론조사를 실시하고 있습니다. 잠시만 응답해주시면 감사하겠습니다. 조사기관 전화번호는 02-333-7223번입니다.

1. 응답자 특성

Q01. 선생님의 연령은 어떻게 되십니까?

- 만 19세 미만(☞ Q99-1.종류) 만 19세 이상 20대 30대
 40대 50대 60대
 70대 이상

Q02. 선생님의 성별은 어떻게 되십니까?

- 남성 여성

Q03. 선생님의 현재 주민등록상 거주지는 어떻게 되십니까?

- 서울 경기, 인천 대전, 충청, 세종 광주, 전남, 전북
 대구, 경북 부산, 울산, 경남 강원, 제주

2. 주요 여론질문

Q04. 선생님께서는 현재 문재인 대통령의 국정수행에 대해 어떻게 평가하십니까?

- 매우 잘하고 있다 대체로 잘하고 있다 대체로 잘못하고 있다
 매우 잘못하고 있다 모르겠다/무응답

Q05. 선생님께서는 신임 조국 법무부 장관의 장관직 수행에 대해 어떻게 생각하십니까?

- 장관직을 계속 수행해야 한다 장관직을 사퇴해야 한다 모르겠다/무응답

3. 성향질문

Q06. 선생님께서는 평소 정치에 얼마나 관심이 있으십니까?

- 많이 있다 약간 있다 별로 없다 전혀 없다
 모르겠다/무응답

Q07. 선생님께서는 내년 4월로 예정된 제21대 국회의원 선거에 투표하실 의향이 있으십니까?

- 반드시 투표할 것이다 투표할 것이다 투표하지 않을 것이다
 절대 투표하지 않을 것이다 모르겠다/무응답

Q08. 선생님의 정치적 이념 성향은 다음 중 어디에 해당한다고 생각하십니까?
 매우 보수적 약간 보수적 중도적 약간 진보적
 매우 진보적 모르겠다/무응답

Q09. 선생님께서는 평소 여론조사에 대해 어떻게 생각하십니까?
 매우 믿을만하다 대체로 믿을만하다 대체로 믿을 수 없다
 절대 믿을 수 없다 모르겠다/무응답

Q10. 선생님께서는 다음 중 어느 정당을 지지 혹은 선호하십니까?
 보기 순서는 무작위입니다. (보기 ~번 로테이션)
 더불어민주당 자유한국당 바른미래당 민주평화당
 정의당 우리공화당 기타 정당 지지정당이 없다
 모르겠다/무응답

4. 배경질문

Q11. 선생님의 현재 직업은 다음 중 무엇입니까?
 사무/기술직, 경영/관리직, 전문/자유직 판매/서비스직, 기능/숙련공, 일반직업직
 자영업 농림어업 가정주부
 학생 무직/은퇴 모르겠다/무응답

Q12. 마지막 질문입니다. 실례지만, 선생님께서는 학교를 어디까지 마치셨습니까? 중퇴는 졸업에 포함
 되지 않습니다.
 고등학교 졸업 이하 대학 재학 / 졸업
 대학원 재학 이상 모르겠다/무응답

99. [맺음말] 설문에 끝까지 응답해주셔서 대단히 감사합니다.

99- 1. [종료처리] 죄송합니다. 선생님께서는 본 여론조사의 조사대상자가 아닙니다. 감사합니다.

99- 2. [활동완료 종료] 죄송합니다. 표본할당이 완료되어 선생님께서는 더 이상 조사가 진행되지 않습니다.
 거듭 죄송합니다.

부록 2. 전화면접 조사용 설문지

중앙여론조사심의위원회 과제 연구 조사

안녕하세요? 저는 여론조사 전문회사인 한국리서치의 면접원 ____입니다.
우리 사회 현안에 대한 국민 여론을 조사하고 있습니다.
잠시만 시간을 내주시면 고맙겠습니다.
본 조사에 관한 문의사항은 02-3014-1000로 연락주세요.

[응답자 선정 질문]

[연령대]

선문1. 선생님의 연세는 만으로 어떻게 되십니까? 만____세

1. 18세 이하 ☞ 면접중단
2. 19~29세
3. 30~39세
4. 40~49세
5. 50~59세
6. 60~69세
7. 70세 이상

[성별]

선문2. 선생님의 성별은 무엇입니까?

1. 남자
2. 여자

[지역]

선문3. 선생님께서는 현재 주민등록상 거주지는 어떻게 되십니까?

- | | | | |
|--------|--------|--------|--------|
| 1. 서울 | 2. 부산 | 3. 대구 | 4. 인천 |
| 5. 광주 | 6. 대전 | 7. 울산 | 8. 경기 |
| 9. 강원 | 10. 충북 | 11. 충남 | 12. 전북 |
| 13. 전남 | 14. 경북 | 15. 경남 | 16. 제주 |
| 17. 세종 | | | |

[여론 문항]

[대통령 국정수행 평가]

문1. 선생님께서는 현재 문재인 대통령의 국정수행에 대해 어떻게 평가하십니까?

1. 매우 잘하고 있다
2. 대체로 잘하고 있다.
3. 대체로 잘못하고 있다.
4. 매우 잘못하고 있다.
9. 모르겠다./무응답

[최근 이슈]

문2. 선생님께서는 신임 조국 법무부 장관의 장관직 수행에 대해 어떻게 생각하십니까?

1. 장관직을 계속 수행하여야 한다.
2. 장관직을 사퇴 하여야 한다.
9. 모르겠다./무응답

[성향 문항]

[정치관심도]

문3. 선생님께서는 평소 정치에 얼마나 관심이 있으십니까?

1. 많이 있다.
2. 약간 있다.
3. 별로 없다.
4. 전혀 없다.
9. 모르겠다./무응답

[투표의향]

문4. 선생님께서는 내년 4월로 예정된 21대 국회의원 선거에 투표하실 의향이 있으십니까?

1. 반드시 투표할 것이다.
2. 투표할 것이다.
3. 투표하지 않을 것이다.
4. 절대 투표하지 않을 것이다.
9. 모르겠다./무응답

[이념성향]

문5. 선생님께서는 정치적 성향은 다음 중 어디에 해당한다고 생각하십니까?

1. 매우 보수적
2. 약간 보수적
3. 중도적
4. 약간 진보적
5. 매우 진보적
9. 모르겠다/무응답

[여론조사 효용성]

문6. 선생님께서는 여론조사에 대해 어떻게 생각하시는지요?

1. 매우 믿을 만하다
2. 대체로 믿을 만하다.
3. 대체로 믿을 수 없다.
4. 절대 믿을 수 없다.
9. 모르겠다/무응답

[정당지지도]

문7. 선생님께서는 다음 중 어느 정당을 지지 혹은 선호하십니까? 보기는 무작위 순서입니다.(보기 1-6 Rotation)

1. 더불어민주당
2. 자유한국당
3. 바른미래당
4. 민주평화당
5. 정의당
6. 우리공화당
7. 기타 정당
8. 지지정당이 없다
9. 모르겠다/무응답

[응답자 특성]

[직업]

배문1. 선생님의 직업은 무엇입니까?

1. 사무/기술직, 경영/관리직, 전문/자유직
2. 판매/서비스직, 기능/숙련공, 일반작업직
3. 자영업
4. 농림어업
5. 가정주부
6. 학생
7. 무직
8. 은퇴
9. 모르겠다./무응답

[학력]

배문2. 실례지만, 선생님께서는 학교를 어디까지 마치셨습니까? 중퇴는 졸업에 포함되지 않습니다.

1. 고등학교 졸업 이하
2. 대학 재학 또는 졸업
3. 대학원 재학 이상
9. 모르겠다./무응답

[응답해 주셔서 감사합니다. 지금까지 여론조사 전문기관 한국리서치였습니다.]
[한국리서치는 서울 강남구 봉은사로 179 H-Tower에 위치해 있습니다]