

연구논문

제 16대 국회의원 선거의 예측조사에 대한 사후적 검증

A Post-Examination of Forecasting Surveys for the 16th General Election

홍 내 리* · 허 명 회**
Nae-Ri Hong · Myung-Hoe Huh

2000년 4월 13일의 제 16대 국회의원 선거에서 전화조사와 출구조사를 활용하여 KBS, SBS, MBC 등 TV 방송사들은 당선자 예측방송을 한 바 있다. 그러나 결과적으로 부정확한 것으로 드러났고 이에 따라 통계조사에 대한 일반인의 불신감만 키우는 결과가 되었다. 이 연구의 목적은 다음과 같다. 1) 4·13 총선거 예측에서 전화조사와 출구조사가 실제로 어떠한가? 2) 무엇이 문제였는가? 3) 무엇을 개선해야 하는가? 등이다. 주요 결론은 전화조사에서 재통화(call back)에 대하여 느슨한 조사규칙이 적용되는 등 확률표집의 원칙이 충실하게 준수되지 않았으며 출구조사에서는 투표소 표집 수가 너무 작았다는 것이다. 또한 근본적인 개선을 위해서는 선거 여론조사에 관한 법의 개정이 필요함을 지적하였다.

On the election day of the 16th Korea National Assembly, nationwide TV stations such as KBS, SBS and MBC broadcasted election forecasting based on telephone surveys and exit polls. The result turned out to be significantly wrong, undermining general audience's confidence on the "scientific" surveys. The purpose of this study consists of the followings. i) What form of telephone surveys and exit polls were adopted by survey institutions in real field? ii) What were the problems? iii) What can we do to make it straight? Major findings of the study include the ignorance of call-back rules in telephone survey and the inadequate number of sampled election posts to make satisfactory forecasting. It is pointed out that it is necessary to amend election laws related to election opinion polls to make a substantial progress.

* 고려대학교 통계학과 석사과정 졸업, e-mail: hongnl@dreamwiz.com

** 고려대학교 정경대학 통계학과 교수. (136-701) 서울특별시 성북구 안암동 5가 1
전화 02-3290-2235, 팩스 02-924-9895, e-mail: stat420@mail.korea.ac.kr

I. 들어가며

선거예측조사는 여론조사의 잘잘못이 확연히 가려질 수 있는 거의 유일한 여론조사이다. 예를 들어, 현 정부의 남북정책에 대한 여론조사에 대해서는 그 결과가 맞았는지 혹은 틀렸는지를 알기 어렵다. 그러나, 선거예측조사에서 주요한 사항인 지지율에 대한 답은 투표가 마감된 후 불과 며칠, 또는 몇 시간만 지나면 밝혀지게 된다. 선거예측조사의 역할 및 필요성에 대하여는 많은 논란이 있지만, 통계적 측면에서는 한 나라 조사연구의 질적 수준을 살펴볼 수 있는 좋은 기회인 것이 사실이다.

본 논문의 목적은 제16대 국회의원 선거(2000년 4월 13일 실시, 이하 4·13 총선거로 약칭) 예측조사를 여러 통계적 관점에서 검토하고 문제점을 파악한 뒤 궁극적으로 선거 여론조사 방법의 개선방안을 찾는 데에 있다. 구체적으로 본 연구는 4·13 총선거에 대한 예측조사로서 KBS, SBS, MBC 등 TV 방송사의 의뢰로 5개 조사회사에 의해 선거운동기간 동안 선거일 당선자 예측보도를 위해 계획되고 실시된 조사만을 대상으로 한다. (KBS와 SBS는 4·13 총선거를 위한 예측조사를 미디어 리서치, 코리아 리서치, 한국 리서치, TN 소프테스 등 4개 조사회사에 의뢰하였고, MBC는 한국갤럽에 단독 의뢰하였다. 조사회사는 전화조사와 출구조사를 혼합·실시하여 TV 방송사에 득표율 예측자료를 제공하였다.)

KBS/SBS와 MBC의 당선자 예측결과를 살펴보면, 227개 선거구 중 KBS/SBS는 21개(9.3%), MBC는 23개(10.1%) 선거구에서 당선자 예측에 실패했다. <표 1>과 <표 2>를 보면, 민주당 후보의 당선을 예상했으나 한나라당 후보의 당선으로 인해 예측 오류가 발생한 경우는 KBS/SBS, MBC가 각각 15건, 14건임을 알 수 있다. 이는 이번 총선거 예측에서 전반적으로 여당인 민주당이 과대 예측된 반면, 한나라당은 과소 예측되었음을 보여주는 것이다.

KBS/SBS는 총 21개 예측 오류 중 15개(71.4%)가 출구조사를 통해 최종 예측치를 산출한 선거구에서 발생했고, MBC의 경우에는 총 23개 중 15개(65.2%)가 그러하였다. KBS/SBS는 총 77개 선거구에서 출구조사를 실시하였고, 그 중 15개 선거구에서 예측 오류가 발생했다. 특히, 이 15

개 예측 오류는 모두 예상과는 달리 한나라당 후보가 당선됨으로써 발생한 것이다. MBC는 총 79개 선거구에서 출구조사를 실시하였고, 그 중 15개 선거구에서 예측 오류가 발생했다. KBS/SBS와 마찬가지로 한나라당 후보가 당선됨으로써 예측 실패로 돌아간 곳이 13개 선거구로 많았다. 앞서 살펴보았던 당선자 예측에서의 한나라당 과소 예측의 경향은 출구조사에서 두드러지게 나타난 현상임을 알 수 있다.

〈그림 1〉은 한나라당과 민주당 후보자에 대한 예측오차를 보여준다. 이에 의하면 KBS/SBS의 전화조사에서 예측오차의 위·아래 사분위수 차이가 6~7%p로서 표준오차로 환산하면 4.4~5.2%p 정도이며 출구조사에서는 2%p 정도의 편향이 존재한다. 정도(精度) 측면에서는 출구조사가 전화조사에 비하여 다소 나은 편이다. MBC 조사도 KBS/SBS와 유사하지만 전화조사에서 한나라당에 대한 과소예측 편향과 출구조사에서 민주당에 대한 과대예측 편향이 뚜렷이 드러난다.

〈표 3〉은 227개 선거구 당선자에 대한 예측오차가 $\pm 5\%$ p 미만의 경우와 이상의 경우를 방송사·조사방법별로 분류해본 것이다. 전체적으로는 예측오차가 $\pm 5\%$ p 미만으로 나타난 상대빈도가 66.5%(KBS/SBS)·63.4%(MBC)에 불과하고 전화조사에 비하면 출구조사가 오히려 정확한 편임을 볼 수 있다. 그러나 100% 가까이 정확성이 기대되는 출구조사에 대한 예측에서도 오류 수가 상당한 것이 사실이기 때문에 전화조사 및 출구조사 등 선거예측 방법이 전반적으로 실제 어떤 문제점을 갖고 있는지 엄밀히

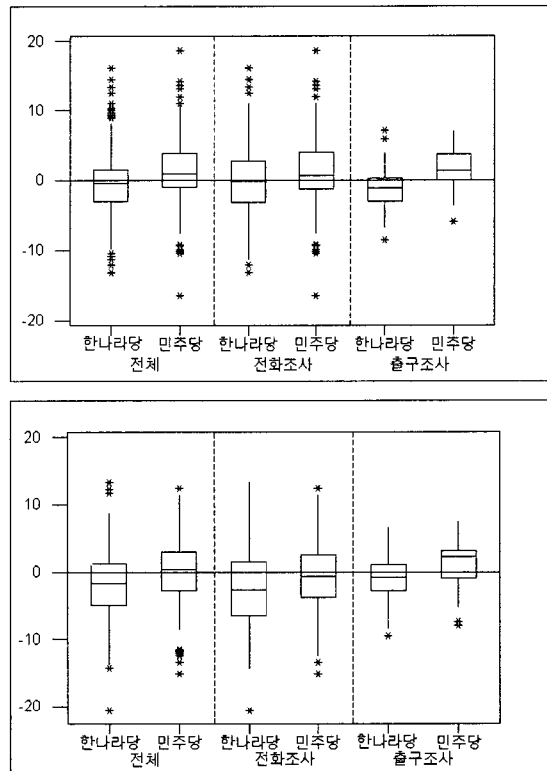
〈표 1〉 KBS/SBS의 당선자 예측결과

예측 \ 결과	한나라당	민주당	자민련	기타	합계
한나라당	94	1	0	0	95
민주당	15	95	0	2	112
자민련	0	0	12	0	12
기타	3	0	0	5	8
합계	112	96	12	7	227

4 조사연구

〈표 2〉 MBC의 당선자 예측결과

예측	결과				
	한나라당	민주당	자민련	기타	합계
한나라당	96	4	0	0	100
민주당	14	91	1	1	107
자민련	0	1	11	0	12
기타	2	0	0	6	8
합계	112	96	12	7	227



〈그림 1〉 KBS/SBS(위), MBC(아래)의 정당별 예측오차

〈표 3〉 전화조사와 출구조사의 당선자 득표율 예측결과 비교

KBS/SBS	±5%p 미만	±5%p 이상	전체
전화조사	83 (55.3%)	67 (44.7%)	150 (100%)
출구조사	68 (88.3%)	9 (11.7%)	77 (100%)
전체	151 (66.5%)	76 (33.5%)	227 (100%)
MBC	±5%p 미만	±5%p 이상	전체
전화조사	77 (52.0%)	71 (48.0%)	148 (100%)
출구조사	67 (84.8%)	12 (15.2%)	79 (100%)
전체	144 (63.4%)	83 (36.6%)	227 (100%)

검토해볼 필요가 있다고 하겠다.

15대 총선거에 대한 검증적 연구에 대하여는 허명희 (1998)의 제 1장이 있으며, 16대 총선거 예측에 대하여는 한국조사연구학회 김정훈 (2000)의 2000년 춘계학술대회 토론과 류제복 (2000)과 박용치 (2000) 등의 선행연구가 이 연구의 기본 골격이 되었음을 밝혀둔다.

II. 선거예측조사의 실제

KBS/SBS와 MBC의 의뢰로 이루어진 5개 조사회사들의 전화조사 및 출구조사가 실제로 어떻게 진행되었는가를 살펴보고 하겠다. 자료 접근 및 담당자 인터뷰에 있어 KBS/SBS 조사에 대하여는 비교적 자유로웠으나 MBC 조사에 대하여는 제한적이었으므로 분석의 폭에 있어 불균형이 초래되는 결과가 되었음을 미리 알려둔다.

1. 전화조사

KBS/SBS(미디어 리서치, 코리아 리서치, 한국 리서치, TN 소프레스)는 경험정도에 따라 2~3회의 전화조사를 실시하였다. 한국 리서치와 TN 소

6 조사연구

프레스는 막판에 출구조사의 필요성을 감지하였으나 미처 준비가 안된 지역에 대해서 투표당일 전화조사를 실시하였다. MBC(한국갤럽)는 선거 운동 기간동안에 지역구별로 2~4회의 전화조사를 실시하였고, 4월 13일과 14일에 두 차례의 사후 전화조사를 실시하였다. 사후조사는 후보 지지 이유, 투표 영향 요인, 이슈에 대한 견해 등 투표 행태에 관한 정보 제공 및 전국구 득표율을 예측하기 위해 실시하였다.

1) 응답자 선정방법

미디어 리서치는 읍/면/동 유권자 비율에 따라 표본가구 수를 정하고, 선거구 내 전화번호들을 지역별로 구분한 뒤 구분된 각 지역 내에서 계통추출로 표본가구를 추출하였다. 응답자의 선정은 해당 읍/면/동의 성, 연령 비율에 따라 할당된 표본 수를 채워나가는 방식으로 하였다. 코리아 리서치는 2000년 1월 최신 전화번호부를 이용하여 미디어 리서치와 비슷한 방법으로 응답자를 선정하였다.

한국 리서치는 읍/면/동 유권자수 비율에 따라 표본가구 수를 정한 후 최근의 전화번호부에서 응답가구를 무작위 추출하였다. 1차, 2차 전화조사에서 응답자 선정은 각 선거구의 성, 연령 비율에 따라 할당된 표본 수를 채워나갔으며, 3차 전화조사에서는 성, 연령, 읍/면/동 비율에 따라 표본 수를 할당하였다. 2개 이상의 행정구역이 하나의 선거구를 구성하는 복합선거구의 경우에는 행정구역별로 별도의 할당추출을 하였다.

TN 소프레스 역시 표본가구는 전화번호부를 이용하여 무작위 추출하였고, 표본가구 내에서 응답자의 선정은 선거구내 읍/면/동, 성에 대한 유권자 분포를 토대로 할당 추출하였다. TN 소프레스는 다른 조사회사들과 달리 연령을 할당 변수로 고려하지 않았다.

한국갤럽은 읍/면/동별 할당 수를 정한 후 각 동(洞)의 담당 조사원들이 정해진 수만큼의 응답을 받아내도록 하였다. 담당 동의 전화번호부에서 전화번호를 추출하는 것과 통화가 된 가구에서 응답자를 선정하는 것 등을 모두 조사원들에게 일임하였다.

2) 재통화(call-back) 규칙

이번 총선거 예측조사에 참여했던 5개 조사회사는 부재중, 통화중, 조사거절의 상황에서 모두 재통화를 하지 않았다. 그 이유는 표집방식이 지역, 성, 연령 등을 할당변수로 하여 할당 표본 수를 채워가는 것이었으므로 재통화를 통해 처음 접촉자에게 응답을 받아내려는 노력이 굳이 필요하지 않은 것으로 보았기 때문이다. 따라서 응답률¹⁾도 매우 낮았을 것으로 예상되는데, 전화조사에서 응답률을 조사한 TN 소프레스에 의하면 응답률은 평균 16.2%였다고 한다. TN 소프레스는 경합이 예상되나 출구 조사를 실시하지 못한 지역에 대해서 선거 당일 투표자 사후조사를 실시하였는데, 이때는 재통화를 하였으며, 원칙적으로 3회까지 시도하였다.

3) 질문지의 구성

미디어 리서치는 가장 먼저 5점 척도로 투표의향을 묻는다. 그리고 나서 바로 지지후보에 대하여 묻되, 응답하지 않은 경우에는 어투를 바꾸어 한 번 더 질문한다. 지지정당이나 인구통계적 변인과 같이 기본적인 사항들을 질문한 후 마지막으로 지지후보에 대해 다시 확인하는 질문을 한다.

매 전화조사마다 질문의 변화가 조금씩 있었지만, 코리아 리서치도 투표의사, 지지후보, 당선가능 후보, 정당기준 선호도, 응답자 특성(성, 연령, 거주지역, 직업, 학력, 소득, 원적)의 순서로 질문하였다.

한국 리서치는 인구통계적 특성(성, 연령, 거주 읍/면/동), 투표권 유무 및 투표의향, 지지후보 및 지지정당, 사회경제적 특성(학력, 직업, 고향) 등을 질문하였다. 지지후보가 없다고 답하는 경우에는 호감 가는 후보를 물었고, 지지후보가 없는 이유(혹은 결정하지 못한 이유)에 대해 질문하였다.

TN 소프레스는 지지후보에 대해서 ‘~ 중 어느 후보가 당선되면 가장

1) 여론조사에서 ‘응답률’은 두 가지 의미로 쓰인다. 무응답률에 반대되는 의미로 쓰이기도 하지만, 거절하지 않고 조사에 응한 비율을 의미하기도 한다. 여기에서는 후자의 의미로 쓰였다.

좋을 것 같습니까?’라는 표현으로 질문하였고, 이에 응답하지 않는 경우 한번 더 생각해볼 것을 부탁하면서 재차 질문하였다. 투표의사에 대해서는 대부분의 유권자가 솔직한 대답을 하지 않기 때문에 투표참여 의향을 직접 묻지 않고 ‘투표할 것 같습니까? 투표하지 않을 것 같습니까?’와 같이 간접적으로 질문하였다. 특히, TN 소프레스는 투표 당일 언제 투표할 것인지를 질문하였다.

한국갤럽은 성, 연령, 학력, 원적, 거주 지역을 기본적으로 질문하고, 지지후보와 이에 무응답인 경우 호감 가는 후보를 물었다. 투표의향은 4점 척도 문항으로 질문하였다.

4) 최종 예측치 산출방법

미디어 리서치는 선거구에 따라 최종 예측치 산출을 위해 사용한 방법이 다르다. 크게는 무응답자를 제외한 방법, 단순판별, 단순누릴, 분류판별, 분류누릴 방법의 5가지로 나눌 수 있다. 판별분석의 경우 지역 상황에 따라 판별변수가 변하는데, 예를 들어 서울에서 한나라당과 민주당의 경합이 예상되는 지역이면 출신지역을 가장 우선적으로 고려하나, 자민련이 우세로 나타나는 지역에서는 출신지역을 판별변수에서 제외하기도 했다. 그리고 각각의 방법에 대해서 전체 응답자를 대상으로 하는 방법이 있고, 투표예상자만을 대상으로, 주요직업에 종사하는 응답자만을 대상으로, 주요지역의 거주자만을 대상으로 하는 네 가지 분석방법이 있다. 따라서 총 20가지의 분석이 가능하다. 여기에 이전 선거의 투표율을 토대로 성, 연령 가중치를 부여한 결과까지 포함하면 모두 40가지 방법에 의한 예상 득표율이 나온다. 이 중에서 그 지역의 판세나 선거에 영향을 미칠 수 있는 여러 요인을 고려하여 가장 타당한 예측치를 최종 예상 득표율로 정하였다.

코리아 리서치는 무응답자 중에서 4점 척도로 조사한 투표의향에서 4점인 ‘반드시 투표하겠다’에 응답한 무응답자만을 대상으로 판별분석을 하였다. 먼저, 투표의향에 따라 추려진 무응답자에게 15대 총선에서의 선거구별, 성별, 연령별 투표율을 토대로 가중치를 부여하였다. 그 후에 성, 연령, 학력, 지역, 원적지를 판별변수로 하여 무응답자 판별을 하였다. 응

답자에 대해서도 역시 ‘반드시 투표하겠다’는 응답자에게만 15대 총선에서의 선거구별, 성별, 연령별 투표율을 토대로 가중치를 적용하였다. 그리고 나서, 앞서와 같은 방법으로 판별된 무응답자의 자료를 응답자 자료와 합하여 최종 예측치를 산출하였다.

한국 리서치는 표본의 성별, 연령별 응답자 분포가 해당 선거구의 유권자 분포와 일치하도록 가중치를 부여하였다. 그런 다음 투표의향자 비율이 60% 이상인 경우에는 지지후보에 대한 1차 질문에서 응답자가 응답한 결과를 기준으로 후보별 예상 득표율을 추정하였다. 그러나, 가중치를 부여한 투표의향자 비율이 60% 미만인 경우에는 투표의향자의 지지후보에 대한 응답과 무응답자의 호감 가는 후보(또는 호감 가는 후보가 없을 경우에는 지지정당)에 대한 응답을 고려한 결과를 기준으로 후보별 예상 득표율을 추정하였다. 투표일 당일 전화조사의 경우에는 투표자가 실제 투표한 후보와 투표할 의향이 있다고 말한 응답자가 지지하는 후보를 조사하였으나, 후보별 예상 득표율 추정을 위해서는 이미 투표를 마친 응답자들의 지지후보만을 고려하였다.

TN 소프레스도 기본적으로 투표의향이 있는 응답자들의 자료만을 포함하였다. 그리고, 전화면접원별로 후보 지지도를 분석하여, 다른 면접원과 상이한 패턴을 나타내는 면접원의 조사가구에 대해서는 추적조사를 하여 부정확한 조사를 행한 것이 밝혀지면, 해당 면접원의 자료는 제거한 후 예측치를 산출하였다.

한국갤럽은 무응답자에 대해 판별분석을 실시하였다. 도시의 경우 원적지, 성, 연령, 학력, 직업을 판별변수로 주로 사용하였으며, 지방의 경우는 읍/면, 성, 연령, 학력을 주로 사용하였다. 지방의 경우, 선거구 내 읍/면에 거주하는 사람들이 후보로 나오기 때문에 각 후보들이 어느 지역 출신이냐에 따라 지지성향이 매우 달라지므로, 이를 가장 중요한 변수로 생각하였다고 한다. 이때 다른 조사회사와의 차이점은 투표의향에 관계없이 모든 응답자들을 최종 예측치 산출에 반영하였다는 것이다.

2. 출구조사

미디어 리서치는 18개, 코리아 리서치는 20개, 한국 리서치는 19개, TN 소프레스는 20개 선거구에서 출구조사를 실시하여, KBS/SBS는 결과적으로 총 77개 선거구에서 출구조사를 실시하였으며, 한국갤럽은 단독으로 79개 선거구에서 출구조사를 실시하였다. 총 65개 지역구에서 약 방송사가 동시에 출구조사를 진행한 것으로 나타났다.

1) 출구조사 대상 지역 선정방법

미디어 리서치는 1~3차에 걸친 전화조사 결과를 토대로 18개 선거구를 최종 확정하였다. 이 때 득표율 차이만을 절대적 기준으로 삼은 것이 아니고, 다음과 같은 사항들을 고려하였다. 즉, 일반적인 지역 정서와 상반된 조사결과가 나타나는가, 1위와 2위의 지지도 차이가 점차 줄어드는 경향을 보이는가, 정당 지지도 혹은 역대 선거결과와는 상이한 지지판세를 보이는가, 무소속이나 군소정당 후보가 높은 지지도를 얻고 있는가, 대규모 아파트 단지 조성으로 유권자 구성과 성향이 바뀌었는가 등을 고려하여 출구조사 대상 지역 선정에 반영하였다.

코리아 리서치와 한국 리서치는 후보자 등록 이전에 방송과 신문 등의 여론조사 결과를 기초로 30-40개 지역을 후보지로 선정하였고, 2차 전화조사 이후 확정하였다.

2) 투표소 표집방법

미디어 리서치의 경우, 추출 투표소 수는 6개를 기본으로 하되 지역구별 경합정도와 전체 투표소 수, 유권자 수, 복합 선거구 등의 특성에 따라 가감하였다. 표집방법은 편의상 '역대 선거 오차의 최소제곱법'으로 부르는 방법이다. 이 방법의 핵심은 가까운 시점의 역대 선거에서 지역구 전체 결과와 각 투표소별 결과의 차이(오차)가 가장 적은 투표소 순으로 추출한다는 것이다. 이 때 각 선거의 성격, 지역구내 특성, 후보 구도 등에 대한 연구분석을 통해 각 선거결과 오차의 단순함이 아닌 가중합을 구하게 되며, 당연히 과거 선거와 이번 선거간에 선거구 조정이 있

거나 해당 투표소 관할 통/리의 변동이 있는지에 대해 사전에 파악하였다. 이번에 역대 선거 결과로 고려한 선거는 14대 총선, 15대 총선 및 대선, 98년 지자체 선거였다.

코리아 리서치는 평균 6~7개의 투표소를 추출하였다. 사전계획 단계에서는 15대 총선 및 대선 등 ‘역대 선거 오차의 최소제공법’을 이용하여 선거구를 대표할 만한 투표소를 선정할 예정이었으나, 실제로는 최소제공법을 통해 선거구를 대표할 만한 대표 읍/면/동을 추출한 후 해당 읍/면/동 내에서는 투표소를 확률비례표집(PPS)하는 2단계 방식으로 수정하였다.

한국 리서치는 확률비례표집(PPS)방법으로 선거구 당 21개의 투표소를 표집하였다. 배열된 투표소 목록에서 최초 출발점은, 최소제공법에 의해 15대 대선과 15대 총선에서 전체 선거구와 후보별 득표율이 가장 비슷한 투표소로 하였다.

TN 소프레스가 투표소 표집을 위해 사용한 방법은 최소편차제공법, strata 설정법, 순환식 조사법 등 3가지이다. 한국갤럽은 유권자 수와 투표소 위치의 분산여부를 고려하여 10개 정도의 투표소를 추출하였다.

3) 응답자 선정방법

미디어 리서치는 투표소별 유권자 수, 오후 4시까지의 예상 투표자수 추정 등을 고려해 5~7명마다 1명을 선정하였다. 코리아 리서치는 투표자 7명마다 1명을 응답자로 선정하였다. 한국 리서치는 매 5번째 투표자를 응답자로 선정하였다. 1개 투표소당 1~2개 지점에서 조사를 실시하므로, 1개 지점에서만 조사할 경우 응답자는 매 5번째 투표자를 추출하고, 2개 지점에서 조사할 경우 매 10번째 투표자를 추출하였다. 그러나, 오후로 접어들면서 투표자 수가 급격히 줄어들어 응답자 수가 너무 작다고 판정되는 선거구에서는 3명마다 1명으로 규칙을 바꾸어 적용하기도 했다고 한다. TN 소프레스는 매 5~6번째 투표자를 응답자로 선정하였다. 한국갤럽도 마찬가지로 방법으로 응답자를 선정하였다.

4) 출구조사 진행과정

모든 조사회사가 공통적으로 투표후보, 성, 연령만을 질문하였다.

미디어 리서치의 경우, 한 조는 5명으로 구성되었고, 조사는 오전 6시부터 오후 4시까지 진행하였다. 면접원은 질문지 수거함과 조사원 신분증, 방송사 협조 공문 사본을 지참하였으며, 조사는 자기 기입식 응답 후 직접 접어서 수거함에 넣는 방법으로 진행하였다. 조사 거절자 수는 집계하지 않았으며, 선정된 투표자가 조사를 거절할 때는 그 투표자를 제외하고 다시 동일 간격만큼 센 후 선정된 사람으로 대체하였다.

코리아 리서치의 조사기간과 기입방식은 미디어 리서치와 동일하다. 선정된 투표자가 응답을 거절하는 경우에는 별도로 표시한 후 다음 7번째 사람을 조사하였다. 투표소별로 조사 거절자의 수는 집계하였으나, 성과 나이는 기록하지 않았다. 조사 거절률은 약 20-25% 였다.

한국 리서치는 한 선거구에 7개 조(한 조는 4명으로 구성)를 배치하였고, 각 조는 3개의 투표소를 시간대별로 이동하면서 조사하여, 한 선거구에서는 총 21개 투표소가 조사되었다. 조사는 응답자가 직접 기입한 후 수거함을 대신한 가방에 넣도록 하는 방식으로 진행하였다. 선정된 투표자가 응답을 거절하면, 다음 5번째 투표자를 면접하였고, 거절자에 대해서는 면접원이 직접 성과 연령을 기입하였다. 이렇게 하여 집계된 조사 거절률은 지역, 성, 연령별로 다르지만, 약 20% 정도였다.

TN 소프레스도 자기 기입 후 수거용 봉투나 가방에 직접 넣도록 하였다. 조사를 거절하는 투표자에게는 성과 연령만이라도 기입해달라고 부탁하여 성과 연령에 대한 응답을 받아내도록 하였다. 이렇게 하여 집계된 조사 거절률은 약 8%였다.

한국갤럽은 2명이 1조를 이루어 한 개 투표소를 담당하였다. 조사는 오전 8시부터 오후 4시까지 진행되었다. 조사원에게는 무선 인터넷폰을 지급하여 응답내용을 듣는 즉시 입력하여 한국갤럽 컴퓨터로 전송되도록 하는 방식을 사용하였다. 거절자에 대한 집계와 조사는 하지 않았다.

5) 최종 예측치 산출방법

미디어 리서치는 투표소별 유권자수 비율에 따라 가중치를 부여하였

다. 코리아 리서치는 조사 거절자 비율을 고려하여 가중치를 부여하였다. 즉, 거절자 비율이 높은 투표소의 결과에는 상대적으로 낮은 가중치를 부여하는 것이다.

한국 리서치는 대부분의 선거구에서는 응답자 자료만을 이용하여 득표율을 계산하였다. 그러나, 전화조사와 비교하여 특이한 지지패턴을 보이고, 초경합 양상을 보이는 2~3개 지역에 대해서는 성, 연령 유권자 비율 가중치를 부여하였다. TN 소프레스의 경우, 출구조사 최종 예측치는 기본적으로 응답자 자료만을 이용하여 산출되며, 선거구에 따라서는 유권자수 비율에 따라 가중치가 적용된 곳이 있다. 한국갤럽은 각 투표소의 유권자 성, 연령 비율을 토대로 가중치를 부여하였다.

Ⅲ. 선거예측조사의 문제점

15대 총선거가 끝난 1996년 4월 11일 오후 6시 정각, 방송 3사는 일제히 ‘신한국당의 압승’을 예측 보도하였다가 큰 망신을 당했다. 그런데 꼭 4년이 지난 4월 13일 16대 총선거에서 비슷한 실수가 되풀이되었다. 일반 시청자들의 눈에는 여당 쪽으로 편향된 부정확한 여론조사로 비치고 말았다. 이런 상황이 벌어진 데에는, 여론조사를 마치 족집게 점쟁이 정도로 비치게 한 보도 방법도 큰 책임이 있지만, 조사 과정과 예측 방법의 곳곳에 문제가 발생할 수 있는 소지가 있음을 부정할 수 없다. 이 절에서는 선거예측조사 과정의 이러한 문제점들에 대해 지적할 것이다.

1. 전화조사에서의 문제점

전화조사는 전국을 거의 동일 시점에서 커버할 수 있으며, 면접조사나 우편조사에 비해 비용도 훨씬 절감할 수 있다. 더욱이 한 장소에서 조사원들을 감독할 수 있고, 조사원들이 장소를 옮길 필요가 없기 때문에 시간적으로도 절약된다. 인터뷰가 끝난 후에 면접결과에 대한 분석이 즉시 가능하므로 조사를 완성하는 데에도 많은 시간을 절약할 수 있다. 이와

같이 전화조사가 갖는 방법론적인 장점과 지난 80년대에 이루어진 높은 전화 보급률로 인해 전화조사는 여론조사를 위한 효과적인 조사방법으로 자리잡았다. 그러나, 국민 생활 패턴이 많은 변화를 보이면서 기존의 방식대로 전화여론조사를 실시하는 것이 많은 문제점을 드러내고 있다.

1) 할당표집으로 인한 선택편향(selection bias)

우리 나라에서 활용되고 있는 전화조사는 대부분이 최종 조사대상자 선정에 할당표집(quota sampling)에 의존하고 있다. 이번 선거예측을 위한 전화조사에서도 전화번호부에서 전화번호를 추출할 때에는 무작위 추출이나 계통 추출방법과 같은 확률표집방법을 사용하였지만, 최종적으로 응답자를 선정할 때에는 성별, 연령별, 지역별로 일정한 수를 할당해서 표집하였다.

할당표집방법의 사용으로 인해 이번 선거예측에서 전화조사가 갖는 문제점은 두 가지 측면으로 정리할 수 있다. 첫째는 표집오차의 계산이 불가능하다는 것이다. 다른 조사방법도 마찬가지겠지만, 전화조사에서 그 결과의 신뢰성은 표집방법에 크게 좌우된다. 주지하다시피 방법론적으로 가장 안정하고 신뢰할 수 있는 표집방법은 무작위 표집(random sampling) 방법이며, 무작위 표집방법의 최대 장점은 확률이론에 근거하여 표집오차의 계산이 가능하다는 것이다 (양승목 · 김현주 · 조성겸, 1991). 전화조사가 과학적인 조사가 되기 위해서는 표집오차의 계산이 가능해야 한다. 모든 전화조사에서 표집오차를 제시하였지만, 최종적인 응답자 선정이 비확률적으로 이루어지고 있기 때문에 사실상 표집오차의 계산은 불가능하며, 제시된 표집오차를 신뢰할 수 없다.

두 번째는 할당변수 이외의 변수들에 대해서는 모집단 분포를 대표하지 못할 수 있다는 것이다. 원래 할당추출은 조사된 사람들과 조사되지 않은 사람들간의 차이를 강력히 억제할 수 있는 변인이 있는 경우 채택할 수 있는 비확률적 표집방법이다 (조성겸, 1997). 후보 지지성향과 관련하여 타당성이 입증된 할당변수를 사용하여야 하지만, 그렇지 못한 경우 할당표집방법은 매우 위험한 결과를 초래할 수 있다. 충분한 조사 기간과 재통화 규칙 하에서 조사를 실시한다면 할당변수 외의 변수들에 대

해서도 어느 정도 왜곡되지 않은 표본을 얻을 수도 있겠지만, 이번 선거 과정에서 전화조사는 대부분 2~3일 동안에 급하게 이루어졌다.

2) 재통화 부재와 응답자 선정에서의 문제점

면접원은 조사 과정에서 각종 문제에 부딪힐 수 있다. 통화중, 부재중, 결번, 상가번호, 자동 응답기 등의 상황에 부딪힐 수 있으며, 통화가 되더라도 비협조적이거나 집에 선거권이 있는 사람이 없거나, 아이가 받는 경우, 질문을 이해하지 못하는 경우가 또한 발생한다. 또, 한 가정에는 선거권이 있는 어른이 보통 한 명 이상 거주하는 경우가 많기 때문에 이들 중 누구를 응답자로 선정하느냐가 중요한 문제이다. 이러한 상황에서 면접원이 마음대로 응답자를 고르게 되면 심각한 선택편향이 생기게 된다. 이에 대한 경계로 미국의 조사회사들은 1988년과 1992년 선거에서 적게는 2, 3회에서 5회 정도까지 재통화를 시도하였으며, Kish Table이나 Birthday Method와 같은 방법으로 응답자를 선정하는 등의 노력을 기울였다 (Voss, Gelman, and King, 1995).

이번 선거예측조사과정에서 5개 조사회사들은 모두 재통화를 실시하지 않았으며, 가구 내 응답자 선정에 있어서도 규칙이 없었다. 그 결과 접근이 용이한 계층이 과대표집되었을 것으로 보인다.

3) 조사 거절과 무응답

전화조사에서 응답률을 집계한 TN 소프레스에 의하면, 평균 응답률은 16.3%였다고 한다. 나머지 83.7%에는 결번, 유권자가 없는 경우, 통화중, 부재중, 조사 거절인 경우가 모두 포함되어 있으므로, 조사 거절률만을 정확히 알 수는 없었다.

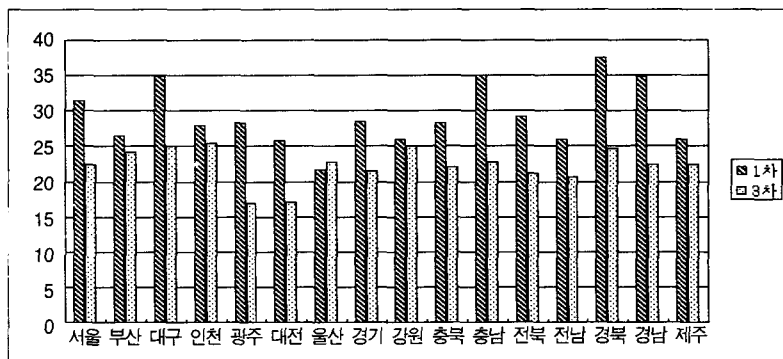
〈표 4〉는 1995년에 리서치 앤 리서치가 우리 나라 선거여론조사에서 면접 성공률이 얼마인가를 알아보기 위해 실험적으로 조사한 결과이다 (노규형·김학량, 1997). 이 실험의 결과를 통해 조사 거절의 정도를 예상할 수 있을 것이다. 조사 성공률이 TN 소프레스가 제시한 것과 비슷함을 확인할 수 있다. 결번과 통화중/부재를 제외하고 통화가 된 1,900가구 중에서는 조사 거절률이 47.4%로 절반에 이르고 있다. 즉, 통화가 된

경우 절반 정도는 조사를 거부한다는 것인데, 이는 매우 심각한 수준이다. 여론조사가 생활화된 미국에서도 처음 전화를 걸었을 때 응답을 거절하는 비율은 줄잡아 60~70%라고 한다 (변상근, 2000년 3월 28일자 중앙일보). 이번 선거과정에서도 전화공해라고 불릴 정도로 많은 여론조사와 선거운동이 전화를 통해 이루어졌고, 자신의 견해를 밝히기 꺼려하는 사회적 태도, 과거 권위주의 시절 피해의식의 잔재로 인한 거부감의 작용으로 많은 유권자들이 조사를 거부했을 것으로 보인다.

〈표 4〉 첫 번째 시도에서의 전화 면접조사 성공률

총통화	결번	통화중/부재	조사 거절	조사 성공
5,668 (100%)	1,029 (18.2%)	2,739 (48.3%)	900 (15.9%)	1,000 (17.6%)

통화가 된 가구 중에서도 절반 가까이는 조사를 거절하게 되는데, 더욱 문제가 되는 것은 조사에 응하는 유권자 중에서 15~35% 정도의 유권자는 지지후보를 묻는 결정적인 문항에 대해 ‘모름’, ‘밝히기 싫음’, ‘아직 결정하지 않았다’와 같이 응답한다는 것이다. 〈그림 2〉는 KBS/SBS의 조사를 수행한 4개 조사회사의 결과를 바탕으로 1차 전화조사와 3차 전



〈그림 2〉 지역별 무응답률 평균

화조사의 지역별 무응답률 평균을 나타낸 것이다. 무응답률이 높게 되면, 무응답자의 지지후보를 무리하게 추측하여 득표율 산출에 반영하게 된다. 그러나, 조사회사마다 무응답자 판정방법이나 무응답자 포함여부가 다를 정도이므로 이에 대한 찬반의견은 분분한 것으로 보인다. 그러므로, 어떤 방법으로 무응답자를 판정하느냐에 따라 그 예측결과는 제각각이게 되고, 이것이 최종 예상 득표율 산출에 상당한 영향을 미쳤을 것으로 보인다.

2. 출구조사에서의 문제점

출구조사를 흔히 ‘투표함에 손을 넣어 조사하는 것과 같은’ 방법이라고 이야기한다. 종전의 유권자를 대상으로 하는 조사에서는 대상자가 실제로 투표를 하는가 하지 않는가 하는 점에서 차이가 생길 가능성이 늘 있었던 것에 비해 출구조사는 투표를 끝낸 사람을 대상으로 하기 때문에 이런 가능성은 거의 없다. 그러나, 실제 조사과정을 살펴보면 계획과 과정상에 많은 문제가 드러나, 출구조사를 통해 정확한 예측을 할 수 있으리라는 기대는 선부른 것이었음을 알 수 있다.

1) 실사와 조사원의 부족

뒤늦게 출구조사 대상 지역으로 선정된 경우에는 실사가 부족하여 투표소의 출입구 위치 파악과 같은 기본적인 사항도 체크하지 못하고 출구조사에 임한 경우가 있었다. 또한, 조사원의 부족도 문제점으로 지적할 수 있다. 한 조사회사의 경우 2명이 1조를 이루어 한 개 투표소를 담당하였다고 한다. 조사원이 부족하면 다음 응답자 선정을 위한 여유가 없기 때문에 조급해지게 되고, 조사 거절자를 설득할 시간적 여유도 없어지며, 거절자에 대한 최소한의 기록(예컨대, 성과 연령)도 할 수 없게 된다.

2) 투표소 수의 부족

이전 선거 결과를 근거로 대표 투표소를 선정하여 출구조사를 진행했

던 미디어 리서치와 코리아 리서치는 6~7개 정도의 투표소를 선정하였다. 한국 리서치는 시간대 별로 투표소를 이동하면서 조사하는 순환식 조사방법을 사용하여 대략 21개 정도의 투표소에서 조사를 하였다. TN 소프레스는 이러한 두 가지 방법을 모두 사용하였는데, 대표 투표소를 선정한 경우 마찬가지로 6~7개 정도의 투표소를 선정하였으며, 순환식으로 진행한 경우에는 20개 정도의 투표소를 선정하였다. 한국갤럽은 한 선거구 당 10개의 투표소를 선정하였다.

여기서는 한 선거구에서 6~7개의 투표소를 추출하는 것이 득표율 예측을 위해 타당한가에 대해서 검토할 것이다. 이를 위해 미디어 리서치와 코리아 리서치 담당 선거구를 대상으로 다음과 같은 모의실험을 실시하였다.

모의실험 1:

- ① 각 선거구에서 조사회사들이 실제로 조사를 실시했던 투표소 수(n)가 얼마인지를 조사하고, 선관위²⁾의 투표소 리스트에서 n 개 투표소를 무작위 추출한다.
- ② 각 후보의 기호를 추출된 투표소에서 획득한 득표수만큼 성성한 후, 이를 규칙성이 없이 배열되도록 섞는다.
- ③ ②의 결과로 마련된 목록에서 7번째³⁾마다 1개씩의 숫자를 뽑아서, 추출된 투표소(i)별로 각 후보(j)의 예상 득표율(p_{ij})을 산출한다.
- ④ 다음과 같은 비추정(ratio estimation) 공식을 이용하여, 해당 선거구 전체의 모비율(각 후보의 실제 득표율)을 추정한다.

2) 중앙선거관리위원회는 전국 227개 선거구에 대해서 투표소별로 각 후보의 득표수와 투표자수, 유권자수, 무효표 수에 대한 정보를 제공하고 있다 (<http://www.nec.go.kr>).

3) 실제로는 각 선거구의 유권자 수를 고려하여 5~6명마다 1명의 투표자를 응답자로 선정하였다. 모의실험에서는 조사 거절이 발생하지 않고, 투표가 이루어진 전시간 동안에 출구조사가 이루어짐을 가정하기 때문에 실제와 같이 응답자를 추출할 경우 표본 수에서 실제와 많은 차이를 보이게 된다.

$$\hat{p}_j = \frac{\sum_{i=1}^n M_i p_{ij}}{\sum_{i=1}^n M_i} \quad M_i : i\text{번째 투표소의 투표자수}$$

- ⑤ ①~④의 과정을 999번 반복하여 평균, 표집오차⁴⁾, 최소값, 2.5% 분위수, 중위수, 97.5% 분위수, 최대값 등의 통계량과 평균 표본 수를 구하였다.

38개 선거구에 대해서 모의실험을 실시한 결과 22개 선거구의 표집오차가 $\pm 5\%p$ 를 넘는 것으로 나타났다. 즉, 이들 선거구에서는 1, 2위간 예상 득표율 차가 10%p라 하더라도 당선자를 확신할 수 없다. 특히, 울산 북구, 경기 양평/가평, 강원 영월/평창, 충남 공주/연기는 표집오차가 $\pm 10\%p$ 를 넘는다. 충남 공주/연기의 경우, 7개의 투표소를 추출하여 출구조사를 여러 차례 실시한다면 임제길(민주당) 후보에 대한 예상 득표율들의 95%는 7.7%와 37.9% 사이에 놓이게 된다. 즉, 임 후보에 대한 예상 득표율의 몬테칼로 분포를 통해 실제 득표율에 대한 95% 신뢰구간이 (7.7%, 37.9%)임을 알 수 있다. 그러나, 이러한 조사 결과는 득표율 예측에 거의 정보를 제공하지 못한다. 5개 투표소를 추출했던 울산 북구도, 노동자들이 주로 거주하는 지역의 투표소가 주로 추출되었다면 실제로 41.8%를 득표한 최용규(민주노동당) 후보가 58.9%를 득표한다는 예측결과가 나올 수도 있다는 것이다. 이런 극단적인 지역을 예로 들지 않더라도 대부분의 선거구가 비교적 큰 표집오차를 보이고 있다. 이를 통해, 허용표집오차를 $\pm 5\%p$ 로 한다는 전제하에, 6~7개의 투표소 또는 응답자 추출을 1/7은 너무 작아서 그 지역 선거구를 제대로 반영하지 못하였다고 말할 수 있다.

3) 투표소 선정방법의 문제

확률비례표집(PPS)방법을 사용한 한국 리서치를 제외하고 대부분의 조사회사는 투표소 추출방법으로 의도적 표본추출(judgemental or purposive

4) $1.96 \times (\hat{p}_j \text{의 표준편차})$

sampling)방법을 사용하였다. 의도적 표본추출법은 통계 이론에 근거하여 표본 추출을 하지 않고, 조사자의 판단에 따라 표본을 추출하는 방법이다. 표본의 주관적 추출 때문에 추출된 표본은 결코 모집단을 확률적으로 대표할 수 없다. 다만, 조사자가 모집단의 속성을 어느 정도 알고있을 때에는 모집단의 속성에 근접하는 표본을 추출할 수 있다 (노규형·김학량, 1997). 이번 선거조사에서는 역대선거에서 선거구의 지지성향과 가장 유사한 투표성향을 보인 투표소를 추출하였다. 그러나, 현재 선거에서의 대결구도가 역대선거와 다른 양상으로 전개된다면 이러한 방법은 오히려 위험할 수 있다. 따라서, 이번 절에서는 이러한 의도적 표본추출방법이 타당하였는지에 대해서 사후적으로 검증하도록 하겠다.

의도적 표본추출방법이 타당하였는가는 추출된 투표소들이 해당 선거구를 제대로 대표하는가를 검증함으로써 밝혀진다. 이를 위해서 다음 모의실험을 실시하였다.

모의실험 2:

- ① 실제로 출구조사를 실시했던 n 개 투표소 하에서 모의실험⁵⁾을 통해 1·2위 후보자의 예상 득표율 q_1, q_2 을 구한다. [→ 의도적 추출방법으로 구한 각 후보 예상 득표율]
- ② 실험 1에서 999번의 반복으로 구한 \hat{p}_1 의 분포에서 q_1 이 어디에 위치하는가를 살펴본다. 2위자에 대해서도 마찬가지로 \hat{p}_2 의 분포에서 q_2 가 어디에 위치하는가를 살펴본다. 이를 위해서 다음을 구한다.

$$\Pr\{\hat{p}_1 > q_1\} = q_1 \text{ 보다 큰 } \hat{p}_1 \text{의 개수}/999$$

$$\Pr\{\hat{p}_2 > q_2\} = q_2 \text{ 보다 큰 } \hat{p}_2 \text{의 개수}/999$$

자료 수집의 어려움으로 실제 진행된 투표소 이름을 알 수 있었던 38개 선거구에 대해서만 모의실험을 실시할 수 있었다. $\Pr\{\hat{p}_1 > q_1\}$ 혹은 $\Pr\{\hat{p}_2 > q_2\}$ 이 0.05이하이거나 0.95이상일 때 선정된 투표소는 1위 혹은

5) 일단 n 개 투표소가 선정된 후 선정된 투표소 하에서 모의실험을 통해 예상 득표율을 구하는 과정은 모의실험 1의 ②~④와 같다.

2위의 지지성향에 있어서 해당 선거구 전체를 제대로 대표하지 못하는 것으로 판단할 때, 모의실험 결과, 38개 선거구 중에서 3개 선거구에서 투표소 선정에 문제가 있었던 것으로 보인다. 즉, 이 3개 선거구의 출구조사 투표소들은 역대 선거에서는 선거구 전체를 대표했는지 몰라도 이번 선거에서는 전체를 대표하지 못했다. 예컨대, 서울 동대문을 선거구의 경우에 $\Pr(\hat{p}_1 > q_1)$ 는 0.9970으로 매우 큰 반면에, $\Pr(\hat{p}_2 > q_2)$ 는 0.0260으로 작음을 볼 수 있다. 즉, 의도적 표집방법으로 투표소를 추출한 결과 김영구(한) 후보(당선자)의 지지자들이 적고, 허인회(민) 후보의 지지자들이 많은 투표소들이 주로 표집되었다는 것이다. 이러한 문제점을 드러낸 선거구는 경기 하남시, 경기 평택을 등이 있다.⁶⁾

4) 조사 거절과 거짓 응답으로 인한 비표집오차의 발생

이번 총선거 예측을 위한 출구조사에서 조사 거절자 수를 집계하고, 거절자에 대한 성과 연령을 기록한 한국 리서치의 자료에 의하면, 조사 거절률은 전체적으로 20%를 약간 웃도는 것으로 나타났다. 성별로 보면, 남성보다는 여성 투표자 층에서, 연령대별로는 남녀 모두 50대 이상 연령층에서 높은 거절률을 보였다. 특히 50대 이상인 여성의 경우, 전체 19개 지역 중에 거절률이 40%를 넘는 지역이 2곳 있었고, 30%를 넘는 지역은 9개나 되었다. 지역별로 보면, 강원, 충북, 충남 등에서 서울이나 인천과 같은 대도시보다 비교적 높은 거절률을 보였다. 높은 거절률은 응답자 선정에 편향을 가져오며, 이는 예측결과를 실제와 멀어지게 한다.

조사 거절 외에 비표집오차의 원인으로 생각할 수 있는 것은 거짓 응답이다. 실제로 출구조사에 투입되었던 조사원들에 의하면, 지방에서는 마을 사람들이 단체로 투표장에 오는 장면이 흔히 목격되었다고 한다. 심지어 마을 선거운동원이 트럭으로 마을 사람들을 실어 나르는 경우도 있었다고 한다. 이런 사람들은 대부분 투표가 끝난 후에도 모여서 함께 가

6) 한 심사위원이 지적한대로 보다 정교한 결과를 얻기 위해서는 모의실험 2의 단계 ①을 여러 번 반복하여 $\Pr(\hat{p}_1 > q_1)$ 과 $\Pr(\hat{p}_2 > q_2)$ 에 대한 평균값을 구하여야 한다. 이 연구에서는 계산적 부담 때문에 반복작업이 수행되지 못하였다.

기 때문에 조사에 어려움이 많다. 우선은 응답자 선정이 어렵게 되고, 주변에 함께 온 사람들이 있기 때문에 응답자는 주변을 의식하게 된다. 이런 경우 거짓 응답이나 무응답이 발생할 가능성이 높아지며, 실제로 지지후보 난에 표시할 때 망설이는 응답자들이 종종 목격된다고 한다.

영국은 1992년 선거예측에서의 실패 이후 패널연구를 통해 조사 거절과 거짓 응답이 예측에 어떠한 영향을 미쳤는가를 사후적으로 분석하였다 (Jowell, Hedges, Lynn, Farrant and Heath, 1993). 우리 나라는 이러한 패널연구가 아직 이루어지지 않아 조사 거절과 거짓 응답의 효과가 어느 정도인지 알 수 없다. 그러나, 조사 거절 및 거짓 응답이 득표율 예측에 미치는 효과가 어느 정도인가를 파악하기 위한 하나의 시도로서 다음과 같은 모의실험을 하였다⁷⁾.

모의실험 3:

- ① 실제로 출구조사를 실시했던 n 개 투표소의 실제 투표자수와 후보별 득표수를 모두 파악한다.
- ② 각 후보의 기호를 추출된 투표소에서 획득한 득표수만큼 생성한 후, 이를 규칙성이 없이 배열되도록 섞는다.
- ③ ②의 결과로 마련된 목록에서 N 번째마다 1개씩의 숫자를 뽑는다.
- ④ ②-③의 과정을 n 개 투표소에 대해 모두 실시한다.
- ⑤ n 개 투표소에서 뽑힌 모든 숫자를 모아서 각 후보별로 단순 득표율을 계산한다.
- ⑥ 위와 같은 실험을 1,000번 실시한다.

실제 조사원들이 오전 6시부터 오후 6시까지 자리를 뜨지 않고 N 번째마다 1명을 선정하는 계통추출의 규칙을 정확히 준수하면서 조사를 진행하였고, 거짓 응답하는 사람과 조사 거절자가 전혀 없었다면, 조사결과 산출되는 예상 득표율은 위 실험의 결과와 비슷할 것이다. 따라서, 전

7) 출구조사 원자료를 제공받을 수 있었던 미디어 리서치 담당 선거구만을 대상으로 모의실험을 하였다.

자의 가정이 충족된다면 예상 득표율과 모의실험결과의 차이는 조사 거절과 거짓 응답에 기인한 것으로 예상할 수 있다. 계통추출의 경우 비교적 정확하게 한 것으로 보이며, 4시 이후의 투표자 수는 그리 많지 않으므로, 전자를 가정하는 것이 그다지 무리는 아닐 것이다.

모의실험결과의 평균(A)을 출구조사에서 실제로 얻는 응답을 근거로 산출한 단순 지지율(B)과 비교하여 두 값의 차이가 큰 지역이 상대적으로 조사 거절 및 거짓 응답이 많이 발생한 지역이라고 할 수 있다. 실험 결과에 의하면, 조사 거절과 거짓응답이 많았던 지역은 울산 북구, 경기 수원 권선구, 경기 동두천/양주, 경북 봉화/울진 등이었다 (각각 B-A = +5.6%p (민주당), +4.1%p (민주당), +4.5%p (민주당), +7.9%p (민주노동당)).⁸⁾ 모두 한나라 후보의 단순지지도가 모의실험 결과들의 평균보다 낮게 나와, 이 지역에서 한나라당 지지자들이 대거 응답을 거절하거나 거짓응답을 하였음을 알 수 있었다. 반면에 상대방 후보에게서는 반대의 현상이 나타났다. 지역별로 조사 거절과 거짓 응답 효과에 차이가 존재하지만, 위 대상 선거구에 대해 살펴본 결과 조사 거절과 거짓 응답으로 인해 득표율 예측에서 대략 2%대의 차이가 생김을 알 수 있었다.

5) 조사방식의 문제

일부 조사회사들은 비밀보장을 통해 좀 더 성실한 응답을 유도하기 위해서 변형된 비밀투표방식을 사용하였다. 응답자로 하여금 스스로 질문지에 기입하게 한 후에 모의투표함의 형식을 빌린 봉투나 가방 같은 것에 직접 넣도록 하였다. 반면에, 다른 조사회사는 조사원이 투표자로부터 응답을 듣는 즉시 단말기를 이용하여 데이터를 본사로 전송하였다. 이렇게 비밀이 전혀 보장되지 않는 방식 하에서는 사회 요망성(social desirability)으로의 편향이 발생할 수 있을 것이다.

〈표 5〉는 이러한 두 방식에 의해 동시에 출구조사가 진행된 부산 북/강

8) 여기서는 크기 500의 단순임의표본이 갖는 신뢰수준 95%에서의 단측 최대 표집오차 3.7%를 임계적 기준으로 하였다. $1.645 \sqrt{0.5 \cdot 0.5 / 500} = 0.0368$, 즉 3.7%. 단순임의추출이 가정된 이유는 조사 투표소가 고정되어 있기 때문이다.

서울 1, 2위 후보의 예상 득표율을 나타낸 것이다. 면접방식에서 민주당 노무현 후보의 득표율이 과대 예측됨을 확인할 수 있다. 이는 한나라당 후보를 지지하는 것이 지역감정에 편승하여 투표를 한 것으로 보일 수 있다는 염려 때문에 한나라당에 표를 던진 투표자들이 조사를 거절하거나 거짓 응답을 했기 때문인 것으로 보인다.⁹⁾ 즉, 지역감정에 의한 후보자 선택은 부당하다라는 사회 요망성이 응답에 영향을 미친 것이다.

〈표 5〉 조사방식에 따른 예상 득표율의 차이

조사방식	후보	
	허태열(한나라당)	노무현(민주당)
자기기입방식(KBS/SBS)	50.4%	37.0%
면접방식(MBC)	45.0%	42.9%
실제 득표율	53.2%	35.7%

6) 300m 거리제한의 문제

이번 출구조사는 투표소로부터 300m 밖에서 이루어졌다. 선관위의 제재가 심하지 않았던 일부 선거구에서는 실제로 300m 안으로 들어가 조사가 진행되기도 했지만, 이런 경우에는 면접원이 심리적 부담을 안기 때문에 응답자에게도 불안정한 응답환경을 제공하게 된다. 결과적으로, 이러한 규정으로 인해 응답자 선정에 편향이 생기게 되고 조사 거절과 거짓 응답의 발생 가능성이 높아지게 되었다.

IV. 선거예측조사의 개선방안

지금까지 제16대 총선거 예측조사의 결과와 조사방법, 그리고 문제점

9) 논리적으로는, 이러한 두 방식의 차이가 조사방식에 의한 것이 아니라, 선정된 투표소의 차이, 즉 두 방식이 진행된 투표소가 일치하지 않음으로써 생긴 것이라는 주장이 가능하다.

에 대해서 살펴보았다. 이러한 선행작업들이 궁극적으로 목표했던 바는 선거예측조사방법의 개선방안 모색이었다. 이 절에서는 전화조사와 출구조사로 나누어 조사방법의 개선방안에 대해서 살펴보고, 올바른 여론조사의 환경과 토대 마련을 위한 언론기관의 역할, 정확한 여론조사 수행과 정착에 걸림돌이 되는 제도에 대해서 지적해 보기로 한다.

1. 전화조사의 개선방안

1) 조사 거절률의 집계

4·13 총선거 예측조사에 참가했던 5개 조사회사 중에서 조사 거절률을 집계한 조사회사는 없었다. 다만, 낮은 응답률이 조사 거절을 포함하여 부재중, 통화중, 결번으로 인해 응답자와 접촉하지 못한 비율이 높았다는 것만을 말해주고 있을 뿐이다. 높은 조사 거절률은 조사시간이 충분하지 못했음을 의미하기도 하고, 응답자를 어렵게 찾는 과정에서 여러 편향이 생길 수 있기 때문에 조사의 질을 판단하는 데에 하나의 기준이 될 수 있다. 그러나, 조사가 긴박하게 이루어질 수밖에 없는 사정 등을 이유로 하여 조사 거절률을 집계, 발표하지 않는다면, 국민의 알권리를 위해 존재하는 여론조사가 국민을 속이는 위험한 결과를 초래할 수 있다. 또한 조사 거절률을 소홀히 대하는 분위기가 만연되어 이에 무감각해 진다면, 더 이상 조사방법의 발전을 기대할 수 없다.

2) 적절한 재통화 규칙의 마련과 준수

선거관련 여론조사의 경우, 조사시간은 대부분 1~2일 정도로 매우 짧고 조사기관간 경합으로 매우 낮은 비용으로 수행되고 있다. 조사시간이 짧은 경우 재통화를 할 시간적 여유가 없기 때문에 결국 조사시간대에 쉽게 접근할 수 있는 계층이 과대표집되게 된다. 그 결과로 조사자료의 질이 떨어지게 되는 것은 당연한 일이다. 조사 거절자는 고연령층이나 교육수준이 낮은 계층인 경우가 많다. 또, 서울대학교 신문연구소에서 실시한 국민생활시간조사 결과에 의하면 전화조사가 주로 이루어지는 저녁 시간의 경우 직업, 학력 등에 따라 재택자 비율이 달라지는 것으로

나타났다(양승목, 김현주, 조성겸, 1991). 즉, 특정 시간대에 부재중인 경우는 특정한 직업이나 학력의 계층일 수 있다는 것이다.

이번 총선거 예측조사에서는 조사회사들이 모두 재통화를 실시하지 않았다. 할당추출에 대한 과신과 시간·비용의 제약을 이유로 재통화를 하는 것이 현실적으로 불가능하다고 이야기하지만, 바로 이 부분이 거둬진 예측실패의 주요 원인일 수 있으므로, 낮은 수준에서라도 재통화 규칙을 세워 지켜나가야 한다. 예를 들어, 부재중인 경우는 1~2시간 후에도 그 대로일 수 있으므로, 충분한 시간 간격을 두고 2회 정도라도 재통화를 시도하고, 조사 거절인 경우는 거둬진 시도가 오히려 더 큰 거부감을 낳을 수 있으므로 숙련된 면접원으로 교체하여 최소한 한 번의 재통화를 시도하는 식으로 조사의 성격과 예산에 맞추어 적절한 재통화 규칙을 세우는 것이 중요하다. 또, 조사기관은 의뢰인의 비용에 조사방편을 맞추기보다는 정확한 조사를 위해서 필요한 사항임을 이해시키는 적극성을 발휘해야 할 것이다.

3) 조사 거절률과 무응답률을 낮추어야

조사 거절률과 무응답률을 낮추는 방법으로 패널조사방법을 생각할 수 있다. 패널을 구성하여 선거기간 동안 여러 차례 조사한다면 거절률과 무응답률을 낮출 수 있을 것이다. 이 때 패널을 구성하는 응답자들의 성향이 일반 유권자들보다 더 정치적으로 변한다거나 선거에 더 관심을 갖게 되는 등 조사가 거둬지면서 오염될 수 있다. 따라서 하나의 패널을 선거기간 내내 조사하는 방법보다는 2~3개 패널을 구성하여 선거기간 중에 교체하여 조사하는 것이 적절한 방법일 것이다.

4) 패널연구를 통한 사후분석

1992년 영국 총선거 예측조사에서 조사기관들은 사상 최악의 예측오류를 범하였다. 모든 조사의 결과는 한결같이 노동당의 승리를 예상하였지만, 결과는 보수당이 노동당을 지지율에서 7.6%나 앞서면서 승리했다. 현대 통계학의 발상지이며, 여론조사의 오랜 역사를 지닌 영국에서 그런 심각한 오류가 발생했음은 놀라운 사실이다. 그러나 그들은 치밀한 사후

분석을 통해 예측오류의 원인이 무엇인가를 소상히 밝혀내었다. 이러한 사후분석이 가능했던 것은 예측조사과정에서 패널조사방법이 사용되었기 때문이다. (Jowell, Hedges, Lynn, Farrant and Heath, 1993).

이번 총선거 예측오류의 원인은 상당부분 높은 무응답률과 조사 거절을, 거짓응답에 기인할 것이다. 그러나, 무응답자와 거절자를 추적하여 실제 어떻게 응답했는가를 조사함으로써 이러한 계층이 미친 영향을 계량적으로 파악하지 않는 이상 이러한 생각은 추측일 뿐이다. 패널을 통한 사후조사의 경험을 누적시킨다면, 부동층과 거절층의 지지성향과 거짓응답자의 속성 등을 파악할 수 있을 것이다. 따라서 단시간에 효과를 기대하기보다는 긴 안목으로 패널조사를 계획한다면, 이후 더욱 정확한 예측을 할 수 있는 토대가 마련될 것이다.

5) Random Digit Dialing 방법의 적용 가능성

이번 총선거 예측조사에서 5개 회사는 전화조사시 모두 전화번호부를 표집틀로 사용하였다. 그러나, 이 방법은 비등재된 번호와 등재를 원하지 않는 가구의 전화번호가 누락되어 있고, 전화가 없는 가구는 처음부터 제외되며, 전화가 많은 가구는 표본으로 뽑힐 확률이 더 높다는 단점을 지니고 있다. 결과적으로, 사생활 침해를 우려하는 사회적 계층(단독가구 및 젊은 세대들)과 이사율이 높은 저소득 인구층이 과소표집될 수 있어, 표본이 특정 계층에 편중될 위험성이 존재한다.

Gallup의 조사에 따르면, 미국의 가구용 전화보급률은 약 95%이고, 이 중에서 비등재된 전화번호가 약 30%를 차지한다고 한다 (<http://www.gallup.com>). 이 정도이면 약 28.5%의 가구를 처음부터 조사대상에서 배제하는 것이기 때문에 전화번호부를 표집틀로 하는 것은 매우 위험한 방법이다. 자료의 미비로 정확한 비율을 알 수는 없었지만, 우리나라의 경우도 비등재된 전화번호의 비율이 점점 늘어나고 있다고 한다. 등재를 원하지 않을 경우, 관할 전화국에 전화만 하면 등재를 쉽게 취소시킬 수 있어 앞으로 이 비율은 점점 더 늘어날 것으로 보인다. 따라서 전화번호부를 이용한 가구추출방법은 해가 거듭될수록 점점 더 많은 취약점을 드러낼 것으로 보인다. 이에 대한 해결책은 임의번호걸기(Random Digit

Dialing, RDD) 방법이다. 임의번호걸기 방법이 조사회사의 여건상 당장 실현될 수는 없을지라도, 대안모색의 일환으로 이 방법의 적용 가능성에 대해서 살펴볼 때라고 생각한다.

2. 출구조사의 개선방안

1) 조사 거절자에 대한 조사

조사 거절자들에게서 특정한 성향이 나타나지 않는다면, 또는 나타나더라도 그것이 지지후보를 결정하는 데에 영향을 미치지 않는다면 조사 거절자는 선거 예측에서 별로 문제가 되지 않는다. 그러나 이전에도 많은 조사를 통해 드러났듯이 조사 거절자들이 특정한 지지성향을 갖고 있기 때문에 문제가 되는 것이다. 거절률이 높으면 그 투표소의 결과는 상대적으로 신뢰성이 떨어지므로 최종 예측치 산출시 이를 고려해야 한다. 또, 비록 성과 나이만을 묻고 있지만, 거절자의 인구통계적인 속성을 조사한다면 사후 분석을 통해 여러 정보를 얻을 수 있으며, 이는 이후 선거예측조사 계획에 도움이 될 수 있다. 이번 총선거 예측조사에서는 일부 조사회사들만이 거절률을 조사하였는데, 앞으로는 더 세심한 조사설계와 면접원 교육을 통해 거절률을 집계하고, 나아가 거절자의 성과 연령도 정확히 조사해야 할 것이다.

2) 응답자의 비밀성을 보장하는 조사방식

1992년 오하이오주 해밀턴 카운티에서는 출구조사의 면접방식과 비밀투표방식을 비교하는 실험을 하였다 (Bishop and Fisher, 1995). 면접방식에서는 조사원이 질문지를 읽어주고, 응답을 대신 기록하였고, 비밀투표방식에서는 응답자가 스스로 질문지를 읽고 응답한 후 "SECRET BALLOT" 이라고 적힌 상자에 직접 넣도록 하였다. 이 두 방법을 30분씩 번갈아 가면서 실시하였고, 지지후보(대통령, 상원의원, 하원의원)와 노인법(노령자를 위한 새로운 과세)에 대한 찬반, 인구적 변인 등을 질문하였다. 그 결과 대통령후보를 묻는 문항에 대해서는 비밀투표방식(1.2%)에 비해 면접방식(9.9%)에서 더 높은 거절/무응답률을 보였다. 즉, 면접방식

의 조사에서는 일부 투표자들이 응답에 상당한 심리적 부담을 느꼈음을 알 수 있다. 주목할만한 결과는 노인법에 대한 응답에서 나타났다. 非투표/거절/무응답과 같이 의사를 밝히기를 꺼리거나 주저하는 비율이 면접 방식에서 더 높게 나타났다 (29.7% 대 36.1%). 그리고, 면접방식일 때보다 비밀투표방식일 때 노인법에 반대하는 비율이 높음을 알 수 있다. 대면접촉에서는 노령자 복지를 위한 세금 징수에 반대한다는 의사를 자신 있게 밝히지 못하지만, 투표방식에서는 비밀이 보장되기 때문에 좀 더 솔직한 응답을 했다고 할 수 있다.

이번 총선거 예측조사에서는 변형된 비밀투표방식이 사용되었다. 4개 조사회사는 응답자의 자기 기입 원칙은 충실히 지켰지만, 선관위의 제재와 이동의 효율성을 위해 대부분 모의 투표함이 아닌 가방이나 봉투를 이용하여 질문지를 수거하였다. 조속한 시일 내에 관련 법규의 개정으로 완전한 모의투표함을 이용한 비밀투표방식이 사용되어야 할 것이다. 나아가 300m 거리 제한 조항의 폐지로 투표소와 가까운 위치에 응답을 위한 간이공간을 마련하여 완전히 비밀을 보장해 준 상태에서 투표자들의 의향을 묻는 방식으로 출구조사가 진행되어야 할 것이다.

3) 철저한 사전준비

출구조사에서는 실시 예상 지역구에 대한 실사를 충분히 하는 것이 중요하다. 실제로 출구조사를 실시할 지역구 수보다 충분히 많은 예상 지역구를 선정해서 해당 지역구의 투표소 위치조사와 답사, 과거 대비 투표소 구성의 변화 여부, 과거 선거결과 자료의 준비 등 기초적인 작업들이 최소한 이루어져야 한다. 그래야 뒤늦게 혼전이 예상되는 지역이더라도 출구조사를 허둥대지 않고 진행할 수 있다.

우리 나라 대부분의 조사회사는 총선거 예측조사와 같은 대규모 조사를 담당할 만큼 충분한 수의 조사원을 평소에 보유하고 있지 못하다. 따라서 조사원의 충분한 확보와 교육이 무엇보다도 중요하다.

4) 적절한 사후가중치 방법 사용

출구조사시 계통추출 규칙을 엄격히 지켜 응답자를 선정하고, 조사 거

절자가 발생할 때마다 거절자의 성과 연령만이라도 꼼꼼하게 기록한다면, 표본을 통해서 적어도 조사가 완료된 시간까지의 투표자 성, 연령 분포를 추정할 수 있다. 이를 토대로 출구조사 응답자에게 가중치를 부여한다면, 이번 예측조사에서와 같이 50대 이상의 여성이 과소표집된 경우 이를 보정할 수 있다.

5) 투표소 표집에 대해서

투표소의 수를 늘리는 것이 후보자 득표율 예측에 얼마나 효과적인가를 살펴보기 위해서 3.2절에서와 같은 방법으로 모의실험을 하여 보았다. 다만, 여기에서는 모의실험 1에서 사용한 투표소 수의 2배를 추출하였다. 그 결과, 서울, 부산, 5개 광역시에서는 울산 북구를 제외하고는 표집오차가 모두 2~3%대로 낮아졌음을 볼 수 있었다. 즉, 투표소를 2배 정도로 늘릴 경우 예측값들이 훨씬 안정적임을 확인할 수 있다. 지방의 경우에도 표집오차가 훨씬 줄어들기는 하지만, 여전히 4%p 이상이 되는 선거구가 절반 정도 된다. 즉, 지방의 경우 투표소별로 지지패턴이 워낙 상이한 선거구가 많아서 투표소 수를 2배로 늘려도 예상 득표율과 실제 득표율의 차이를 좀처럼 줄이지 못함을 알 수 있다. 따라서 대도시를 중심으로 출구조사 투표소 수를 더욱 늘리는 것이 바람직하겠지만, 지방의 경우에는 소지역적 특성을 고려한 새로운 투표소 추출방안이 계속 모색되어야 할 것이다.

다음으로, 투표소 수는 2배로 유지하되 14명마다 한 명을 계통추출하여 역시 실험 1과 같은 방법으로 모의실험을 하여 보았다. 지면상의 제약 때문에 여기서 그 결과를 모두 보일 수는 없지만, 모든 선거구에서 실제 투표소 수를 유지하고 7명당 한 명을 추출한 경우에 비해서 투표소 수를 2배로 늘리고 14명당 한 명을 추출한 경우에서 표집오차가 줄어들음을 확인하였다. 즉, 전체 표본 수가 비슷함에도 예측값들이 더 안정적임을 나타내었다. 물론 전체 표본 수가 같다고 총 소요비용이 같은 것은 아닐 것이다. 비용의 문제로 인해 투표소 수를 늘리는 것이 현실적 제약을 받는다면, 투표소 수를 적절한 수준으로 늘리되 각 투표소 내의 응답자 수를 줄이는 것이 비용을 크게 증가시키지 않고도 표집오차를 줄

일 수 있는 적절한 방법이 될 것이다. 투표소 수는 늘리고 각 투표소 내 응답자 수를 줄이는 방법으로는 한 조가 한 투표소에서 머무는 시간을 줄이고 여러 개의 투표소를 이동하게 하는 순환식 방법이 있을 것이다.

3. 보도관행의 개선

1) 올바른 개표방송의 정착

방송사들은 이번 총선거 개표보도에서 각자가 실시한 출구조사 결과를 최대한 활용하며 대규모 속보 전쟁을 벌인 바 있다. 그러나, 한 눈에 보아도 오차 범위에 들어있는 예측불허의 경합지역임에도 불구하고, 방송사들 간의 시청률 경쟁 때문에 1위와 2위를 결정지어 보도하는 것이 개표방송의 정형으로 굳어지고 있다. 이렇게 당선자 예측을 위주로 개표방송이 진행되고, 이러한 목적에 맞추어 선거예측조사가 진행된다면, 엄청난 비용을 들여 실시하는 선거예측조사의 자료는 6시 정각 당선자 예측 보도 이후 곧 쓰레기가 되고 마는 것이다.

미국의 출구조사에서는 투표후보만을 묻는 것이 아니고, 응답자의 성, 인종, 학력, 나이, 종교, 이념, 소득수준 등을 자세히 묻고 있다. 여기다 특정 후보에게 왜 투표했으며, 지난 선거에서는 어느 후보를 찍었고, 4년 전과 비교한 미국과 주의 경제상황과 투표자의 개인소득변화 등 국가적이거나 주의 경제상황과 투표자의 개인소득변화 등 국가적이거나 주의 주요현안에 대한 반응까지도 묻고 있다. 이처럼 질문항목이 많은 것은 출구조사의 주목적이 당선자 예측이 아니라 투표자 성향을 자세히 분석해 선거 이후 정당의 주요정책결정에 도움을 주려는 데 있기 때문이다. 실제로 개표가 끝난 다음날 뉴욕타임스 등 주요 신문들은 출구조사를 통해 드러난 유권자들의 정치관과 행태를 상세하게 공개했다고 한다.

시청자들의 관심이 주로 누가 당선될 것인가에 있기 때문에 당선자 예측보도라는 이제까지의 목적을 주로 고려하지 않을 수는 없지만, 결과적으로 개표방송은 조사과정에서 수집한 응답자들의 인구통계적, 사회경제적 속성을 바탕으로 유권자들의 정치행태를 분석하고 이를 정책결정에 반영하게 하여 정치 선진화를 촉진시키는 촉매제로서의 역할을 해야 한

다.

2) 응답률의 공개

공정선거 및 선거부정방지법 제108조 ④항은 ‘누구든지 선거에 관한 여론조사의 결과를 공표 또는 보도하는 때에는 조사 의뢰자와 조사기관·단체명, 피조사자의 선정방법, 표본의 크기, 조사지역·일시·방법, 표본오차율, 응답률, 질문내용 등을 함께 공표 또는 보도하여야 하며...’라고 규정하고 있다. 여론조사결과를 발표하는 대부분의 신문과 방송은 이 조항에서 제시한 대부분을 보도하고 있다. 그러나, 유독 잘 지켜지지 않는 것이 응답률에 관한 것이다. 표본의 크기는 자신있게 보도하지만, 그 만큼의 표본으로부터 응답을 얻기 위해서 몇 명에게 접촉을 시도하였는지는 공개하지 않는다. 처음 선정된 사람에게서 응답을 얻어내지 못하고 실패를 거듭하면서 겨우 응답을 얻어낸 경우, 그 과정에서 표본의 심각한 왜곡이 일어날 수 있다. 또, Overnight-Poll과 같이 단시간 내에 재통화 규칙없이 진행된 조사의 경우 응답률이 낮기 쉽다. 이렇게 응답률은 조사의 질을 판단하는 잣대가 되기 때문에 여론조사결과를 투표후보결정을 위한 정보원으로 사용하는 대부분의 국민들에게는 응답률이 반드시 공개되어야 한다.

3) 표집오차에 대한 정확한 보도

이번 선거과정에서 실시된 모든 전화조사와 대부분의 출구조사는 엄밀히 말해서 확률추출방법에 근거하여 최종 응답자를 선정하지 않았다. 전화조사에서는 가구 내 응답자 선정에서 할당추출이 이루어졌으며, 출구조사의 투표소 추출과정은 확률추출에 근거하지 않았기 때문에 확률이론에 근거하여 표집오차를 계산할 수 없다. 그럼에도 불구하고, 신문과 방송은 항상 표집오차를 보도한다. 게다가 균집추출인 경우의 표집오차는 무작위추출을 가정하여 보도되는 표집오차보다 크다.

이와 같이 비확률적 표집방법에 의한 조사 결과를 보도할 때에는 표집오차의 계산이 어려우므로, 예측보도에 보다 신중할 필요가 있다. 또한, 표집오차에 대한 이해를 돕지는 못하더라도, 이를 잘못 보도하여 국민들

이 여론조사결과를 이해하는 데에 잘못된 인식을 심어주지 않도록 노력해야 할 것이다.

4. 제도적 측면에서의 개선방안

1) 공정선거 및 선거부정방지법 제167조의 개정

공정선거 및 선거부정방지법(이후 공선법) 제167조 ②항은 ‘텔레비전 방송국·라디오방송국·정기 간행물의 등록 등에 관한 법률에 의한 일간신문사가 선거의 결과를 예상하기 위하여 선거일에 투표소로부터 300미터 밖에서 투표의 비밀이 침해되지 않는 방법으로 질문하는 경우에는 그러하지 아니하며 (즉, 선거인에게 투표한 후보자 상명이나 정당명을 질문할 수 있으며) 이 경우 투표마감시각까지 그 경위와 결과를 공표할 수 없다’라고 규정하고 있다.

투표소에서 나오는 출구가 2개 이상인 경우, 자동차를 이용해 바로 퇴장하는 경우, 2개의 투표소가 인접해있는 경우에 면접원은 300m라는 제한된 거리에서 정확하게 응답자를 선정하기가 매우 어렵다. 투표자를 놓치기 쉽고, 정확한 카운트를 할 수 없으며, 심지어는 어느 투표소에서 투표한 사람인지조차 구분할 수 없기 때문이다.

이 조항에 대한 문제는 법률적인 문제로 정치권에서 해결되어야 할 문제이지만, 선거예측조사가 사회에 미치는 파장을 고려한다면 방송사, 조사회사, 학계, 중앙선거관리위원회 등이 입법청원 등 이에 대해 보다 적극적인 대처를 해야 한다.

2) 공정선거 및 선거부정방지법 제108조의 개정

공선법 제108조 ①항은 ‘누구든지 선거기간 개시일부터 선거일의 투표 마감시각까지 선거에 관하여 정당에 대한 지지도나 당선인을 예상하게 하는 여론조사의 경위와 그 결과를 공표하거나 인용하여 보도할 수 없다’라고 규정하고 있다. 이 조항의 취지는 왜곡된 여론조사를 통한 흑색 선전과 조사결과의 불법이용을 배제함으로써 이러한 조사들이 선거인의 투표에 영향을 미치는 것을 방지하자는 것이다.

그러나, 여론조사결과 사전발표 금지는 헌법상에 명시된 국민의 알권에 대한 침해이자 다양하고 정확한 정보를 바탕으로 국민의 대표를 뽑을 권리에 대한 침해라고도 볼 수 있기 때문에 이에 대한 논의가 필요하다고 본다.

5. 조사현업의 여건 개선

궁극적으로 통계적 선거예측이 제대로 되려면 조사현업의 여건이 개선되어야 한다. 조사업계의 과당경쟁과 최저가 입찰제·적격심사 낙찰제 하에서는 더 이상의 조사업의 발전은 기대하기 어렵다는 한국갤럽연구소의 박무익(2000)의 주장에 동의하지 않을 수 없다. 우리 사회가 조사의 질에 대한 차별적 평가를 지금과 같이 방치하는 한 더 이상의 발전은 기대하기 어려울 것이다. 이 글을 매듭지으며 질적 조사에 대한 사회의 관심을 재삼 촉구한다.

참고문헌

- 김정훈. 2000. 한국조사연구학회 2000년도 춘계학술대회 토론 내용.
노규형, 김학량. 1997. 《선거와 여론조사》 서울: 나남출판사. 25-26, 116-117.
류제복. 2000. “선거예측조사의 신뢰성 증진방안-16대 총선을 중심으로” 《조사연구》 1권 2호, 15-34.
미디어리서치, 코리아리서치, TN소프레스, 한국리서치. 2000. 《(사)한국마케팅여론조사협회 5월 월례세미나 자료집》
박무익. 2000. “우리나라 조사산업의 현황과 전망,” 《조사연구》 1권 2호, 161-177.
박용치. 2000. “선거예측에서 편협의 감소: 거짓 응답을 중심으로” 《조사연구》 1권 2호, 35-57.
양승목, 김현주, 조성겸. 1991. “전화조사에서 가구 내 응답자의 무선험

를표집에 관한 연구,” 《신문학보》 제26호, 189-214.

조성겸. 1997. “대통령선거 여론조사와 할당표집 방법의 문제점,” 《언론과 사회》 제18호, 29-53.

한국갤럽. 2000. “제16대 국회의원 선거 투표행태”

허명희. 1998. 《사회여론조사 - 통계적 연구사례》 서울: 자유아카데미. 52-67.

Bishop, G. F. and Fisher, B. S. 1995. “Secret ballots and self-reports in an exit-poll experiment,” *Public Opinion Quarterly* 59: 568-588.

Jowell, R., Hedges, B., Lynn, P., Farrant, G. and Heath, A. 1993. “The 1992 British election: The failure of the polls,” *Public Opinion Quarterly* 57: 238-263.

Traugott, M.W. and Price, V. 1992. “Exit polls in the 1989 Virginia gubernatorial race: Where did they go wrong?,” *Public Opinion Quarterly* 56: 245-253.

Voss, D.S., Gelman, A. and King, G. 1995. “Preelection survey methodology: Details from eight polling organizations, 1988 and 1992,” *Public Opinion Quarterly* 59: 98-132.