

출구조사(Exit-poll)의 예측력 검토 -16대 대선을 중심으로-

김정훈*

1. 서론

선거예측조사는 실제의 모수치와 표본조사 통계량을 통한 추정치 간의 차이를 확인할 수 있는 유일한 조사라고 해도 과언이 아니다. 따라서 선거예측의 결과는 조사 자체 및 조사기관의 정확성에 대한 일방적 평가 기준으로 인식되고 작용해 왔기 때문에 국내외의 여러 여론조사 기관들은 선거예측조사의 정확성을 높이기 위해 많은 노력을 기울이고 있다. 이러한 노력 중의 하나로 출구조사가 도입되었다. 선거예측에서 전화조사로 예측하는 것은 줄곧 몇 가지 문제점이 노출되고 지적되어 왔는데 출구조사가 그러한 문제점을 보완해 줄 수 있기 때문이다.

전화조사의 문제점은 첫 번째로 할당추출에 따른 편향이다. 표본추출은 가능한 한 확률추출을 해야겠지만, 현재 여러 가지 이유 때문에 대부분 비확률추출로 전화조사가 이루어지고 있다. 대표적인 비확률추출인 할당표집에서는 면접원에 의한 선정오차가 발생하는 문제점이 있다(March and Scarbrough 1990). 두 번째로는 응답자의 재택률에 따른 편향의 발생이다. 전화조사의 경우 재택자만 접촉할 수 있기 때문에 재택률 차이로 인한 표본편파가 발생한다(홍내리, 허명희 2001). 세 번째로 거짓응답, 응답거부 등 응답자의 불성실/비협조로 인해 조사결과가 왜곡되어 선거예측에 어려움을 줄 수 있다. 지난 대선당시 한 후보쪽 참모가 역선택을 우려한다는 주장을 펴으로써 자연스럽게 역선택 개념이 널리 소개되었고 그럼으로써 그럴 가능성을 더 높힌 것이 아닌가하는 논란마저 있었다. 그 외에도 선거에 임박해 중대한 상황변화 요인이 발생할 경우에도 전화조사로 모수치를 추정하는 데는 다소 한계를 가질 수 밖에 없다.(류제복, 2003). 정몽준 대표가 선거 전날 태도변화를 선언한 것은 그 이전의 전화조사 결과를 무의미하게 만들 정도로 큰 사건이었다. 그러나 전화조사 결과로는 그러한 지지철회의 크기 및 방향에 대해 제대로 추정하기가 어려웠다. 이러한 전화조사 예측의 여러 가지 한계점을 극복하기 위해 우리나라에는 투표자 출

1) 미디어리서치 대표

구조사가 지난 1999년 3. 30재보궐선거 때부터 처음으로 도입되었다.

출구조사가 도입된 이후 전화조사만으로 예측을 한 경우와 출구조사로 예측한 경우의 당선자 예측오차를 비교해 보면, 출구조사를 통한 예측의 정확성이 보다 향상된 것을 알 수 있다. 미디어리서치 자체자료를 근거로 살펴보면, 출구조사가 도입된 이후 미디어리서치가 수행한 총 6회의 조사에서 출구조사로 예측을 한 지역의 당선자 오차는 평균 1.6%였으나, 전화조사로만 예측을 한 지역의 당선자 오차는 평균 3.8%로서 출구조사가 보다 더 정확했음을 보여주고 있다(<표 1>).

<표 1> 전화조사와 출구조사의 당선자 예측오차 비교

구분	전화조사	출구조사
6개 전체*	3.8	1.6
16대 총선	3.7	1.7
3회 지방선거	4.3	2.0

주) * 전체 - 99년 3.30보선, 99년 6.3보선, 00년 16대 총선, 02년 3회 지방선거, 02년 8.8재보선, 02년 16대 대선

따라서 선거예측의 정확성 확보를 위해서 현실적으로 출구조사의 수행은 필수적이라 하겠다. 앞으로의 선거에서도 출구조사가 계속적으로 실시될 것이 예상되는 만큼, 과거 수행해왔던 출구조사의 방법을 검토해보고, 예측력을 좌우하는 요인을 살펴봄으로써 보다 정확한 출구조사 수행을 위한 방향을 모색해보는 것이 본 글의 목적이다.

II. 출구조사의 수행역사

국내의 투표자 대상 출구조사는 앞선 언급대로 1999년 3. 30 재보선부터 도입되었는데, 당시 서울 구로구 갑 등 3개 선거구의 20개 투표소에서 출구조사가 실시되었다. 1996년 제15대 국회의원 선거에 이어 1998년 7. 21 재보궐 선거에서 당선자 예측이 어긋난 곳이나옴으로써 선거예측조사 무용론 등 신랄한 비난을 받았던 시점에, 보다 예측력을 높일 수 있는 대안적 방법으로서 출구조사의 도입을 검토, 방송사 등 관련 당사자들의 지대한 관심 속에서 출구조사가 수행되었다. 결과는 당선자 예측오차 평균이 1.3%로 한 곳에서도 틀린 곳이 없이 정확한 예측을 이끌어내었다.

이어서 그 해 6월 3일 이회창 야당총재의 출마로 관심을 모았던 서울 송파갑과 인천 계양/강화갑 2개 지역구의 6. 3보선에서도 출구조사를 실시해 평균오차 2.0%로 정확한 예측을 해냈었다. 1998년의 4. 2보선, 7. 21보선에서의 전화조사를 통한 당선자 예측 평균오차

가 3.8%~9%대인 것에 비해서는 정확성이 크게 향상되었으므로, 업계와 학계, 방송사 등을 중심으로 출구조사의 예측력을 높게 평가하게 되었고, 그 중요성을 크게 인식하게 되었다(<표 2>).

<표 2> 초기의 출구조사 결과(1999년 3. 30 / 6. 3)

선거 구분		투표소 수		실제결과 - 예측결과		당선자 예측오차
		전체	출구조사	공동여당	한나라당	
3.30	구로구乙	48개	5개	한광옥 53.5 - 56.0	조은희 39.4 - 35.4	2.5
	시흥시	65개	5개	김의재 52.5 - 51.9	장경우 47.5 - 48.1	0.6
	안양시장	158개	10개	이준형 45.1 - 45.8	신중대 54.9 - 54.2	0.7
6.3	송파구甲	50개	15개	김희완 36.9 - 34.9	이회창 61.5 - 63.5	2.0
	계양/강화甲	71개	15개	송영길 41.4 - 44.8	안상수 54.4 - 52.3	2.1

주) 이 결과는 KBS의 의뢰로 수행되어 선거당일 투표마감 직후 개표방송에서 예측 보도됨.

두 번의 재보선에서의 성공을 기반으로 2000년 16대 국회의원 총선에서 전국 단위의 출구조사가 기획되어 실시되었다. 출구조사가 본격적으로 실시된 것이 이 때부터라고도 볼 수 있다. 왜냐하면 이전에 실시한 출구조사는 수도권 지역에서만 제한된 몇 개의 선거구를 대상으로 수행된 데 반해 총선에서는 전국단위로 면접원을 선발하여 대규모로 실시하였기 때문이다. 결과는 21개~23개의 지역구에서 당선자 예측이 빗나감으로써 또다시 언론 등에서 선거예측조사에 대해 엄청난 비난 여론이 쏟아졌으며, 출구조사의 신뢰성까지 의심을 받게 되었다. 당시 최초의 전국적인 대규모 출구조사였으므로 2,500여명에 달하는 대규모 조사원 동원과 교육훈련 미흡, 출구조사 대상 지역구 선정의 정교함 부족, 출구조사 투표소추출 방법(sampling)의 미정립, 그리고 조사 진행 현장에서의 여러 문제점 등으로 인해 상대적으로 많은 예측 오류를 보였었다(한국방송공사, 2000). 또한 3표, 11표 등 실제결과가 20표미만의 차이인 4곳을 포함하여 1,000표 차이 미만으로 당락이 결정된 곳이 모두 15개 지역구로 나타나는 등 초접전인 지역구가 많았던 것도 통계적 방법으로는 당선자 예측을 어렵게 만드는 요인으로 작용했었다.

그러나 그 당시의 비난 여론에 가려 정당한 평가를 받지는 못했지만 통계적 정확성의 측면에서는 분명 향상이 있었다. 15대 총선에서는 당선자 예측이 뒤바뀐 곳이 39곳, 당선자예측 평균오차가 6.1%였던 것이 16대 총선에서는 당선자예측 실패 지역구 수의 감소와 함께 당선자예측 평균오차가 4.4%로 줄어들었기 때문이다.

16대 총선이후 2년 만에 실시된 2002년의 제3회 지방선거에서도 방송 3사가 16개 시/도

의 광역단체장 선거예측조사를 실시하였는데, MBC는 전화조사로 예측을 하였으며, KBS와 SBS는 각각 집전/경합 지역에 한해 출구조사를 실시하고 나머지는 전화조사로 예측을 하였다. 결과는 전화조사로 예측한 MBC의 제주도지사 결과만이 예측이 빗나간 것 외에는 방송 3사가 모든 지역에서 당선자 예측을 적중하였으며, 당선자 평균 예측오차도 3.1%로 나타나는 등 선거예측의 신뢰도를 다시 회복시키는 계기가 되었다.

이후 미니 총선으로 불리운 2002년 8월 8일 13개 지역구의 재보궐선거에서도 상대적으로 극히 낮은 투표율로 선거예측이 어렵다는 일부의 주장이 있었음에도, 보다 향상되고 정교한 표본추출방법 등으로 준비된 출구조사를 통하여 모든 지역구의 당선자를 정확히 예측하였음은 물론, 그 평균 예측오차도 2.6%에 그쳐 출구조사를 통한 성공적인 선거예측이 이루어졌었다.

이어서 작년 연말의 16대 대선을 맞아 방송 3사가 당선자의 정확한 예측뿐만 아니라, 통계적 표본 허용오차 범위 내에서 완벽한 득표율 예측을 해냄으로써, 출구조사의 위력을 다시 한번 더 확인시키는 쾌거를 이루어 내었다.

미국, 영국, 프랑스, 호주 등 외국 선진국에서도 선거결과 예측을 위해 출구조사를 실시하고 있지만 이들 선진국에서도 잘못된 예측으로 비난을 받는 경우가 더 많았다. 대표적으로 미국의 ABC, CBS, NBC등 여러 방송국들이 컨소시엄으로 운영하는 VNS(Voter News Service)사가 2000년 대선 당시 플로리다주의 당선자를 번복하는 소동을 겪었으며, 2002년 중간선거에서는 컴퓨터 시스템에 문제가 생겨 조사결과를 발표하지 못하는 일이 발생해 현재 해체될 위기에 처해있는 것으로 알려져 있다.

III. 출구조사의 방법론

1. 표본추출 단계

출구조사의 정확성을 좌우하는 주요요인을 크게 표본오차와 관련된 표본추출 단계, 비표본오차와 관련된 자료수집 단계 두 부분으로 나누어서 살펴보도록 하겠다. 표본추출은 대표성 있는 조사투표소 추출과 조사응답자 추출 두 단계로 구성된다. 투표소 추출 방법으로는 대표구 추출, PPS(Probability Proportional to size) 추출, 단순무작위추출, 층화추출, 집락추출 등 다양한 방식이 있다. 3. 30 재보선, 6. 3 재선거, 16대 총선 등 지역구 단위의 국회의원 선거에서는 역대선거 오차의 최소화승법을 이용한 대표구 추출방식이 주로 사용되었다. 이 방법은 유권자 유입과 유출 및 선거구 조정 등에 따른 급격한 인구변동이

적어야 효과적인 방법이다. 국회의원 선거와 같이 선거구가 좁은 선거에서는 투표소별로 추적을 하여 인구이동, 선거구 조정 등의 문제가 있는 곳은 제외하고 해당 선거구를 대표하는 투표소 선정이 가능하지만 지방선거, 대통령 선거 등 선거구가 넓은 선거에서는 이 방식을 적용하는 것이 문제가 될 수 있다. 광범위한 투표구를 모두 추적한다는 것이 다소 무리가 될 수 있기 때문이다. 따라서 지역구가 넓은 선거에서는 행정구역과 유권자 성향 등을 기초로 몇 차례 층화를 한 뒤 투표소를 추출하는 방법을 주로 이용한다. 지금까지 대통령 선거, 국회의원 선거, 지방선거, 보궐선거 등 모든 선거에서 출구조사가 실시되었는데 그 때마다 투표소 추출방법이 달랐다. 그 동안의 선행 연구나 실행의 사례가 부족했으므로, 여러 가지 방법을 적용해 봄으로써 보다 나은 방법을 찾아보고자 하는 시도였다. 그러나 한두 번 정도 적용한 결과를 갖고서 어느 방법이 가장 타당하다는 결론을 내리기는 어렵다. 보다 많은 방법이 시도되고, 그리고 많은 결과물이 축적될 경우에 이론적으로 뿐만 아니라 현실 경험적으로도 어떠한 방법이 가장 적합하다는 나름의 결론에 도달할 수 있을 것으로 보여 진다.

<표 3> 투표소 추출에 따른 서울시의 16대대선 결과

	노무현	이회창	당선자 오차	1-2위 격차
실제결과	51.3	45.0		
층화추출	51.3	45.0	0.0	0.0
PPS	52.6	43.6	1.3	2.7
대표구	52.1	44.1	0.8	1.7
단순무작위(30회 반복평균)	50.8	45.0	0.5	0.5

투표소가 아무리 대표성 있게 잘 추출되었더라도 추출된 투표소에서의 조사대상 응답자 선정에서의 대표성 확보 또한 중요한 문제이다. 지금까지 출구조사에서 할당추출을 적용한 사례가 있다고는 하지만, 아무래도 할당추출은 적절치 못하다는 생각이다. 철저하게 확률추출이 담보되어야 하고, 현실적으로 그 확률추출을 가장 잘 보장하는 방법으로는 투표소 출구에서 매 k번째 일정 간격마다의 응답자를 선정하는 체계적인 계통추출법이 경험을 통해 가장 적합한 것으로 인식되고 있다.

2. 자료수집 단계

투표소 및 응답자가 통계적으로 아무리 완벽하게 추출되었다고 하더라도 실제의 자료수집 과정에서 원칙준수의 실사진행과 조직적인 관리가 잘못될 경우에는 비표본 오차의 증

가로 인해 결과의 정확성이 크게 훼손될 수가 있다. 추출된 투표소 현장이 실시하기에 문제가 없는지를 검토해보아야 하며, 면접원 모집과 교육, 현장에서의 원칙준수 관리/감독이 철저해야 비표본 오차를 줄일 수 있다. 면접원 교육은 이론 교육뿐 아니라 Role-playing 등 실습교육을 병행함으로써 실제 조사상황에 대한 대처능력을 키워주는 것이 필요하다. 응답을 받는 과정에서는 응답자의 비밀을 철저히 보장하는 것이 아주 중요하다. 그리고 출구조사 감독관이 조사현장을 순회하면서 원칙준수 여부 등을 감독하고 문제 발생시의 시정조치와 함께 현장을 평가하는 책무를 지닌다. 응답자 선정에 대한 계수의 정확성, 계수 위치 및 조사 위치의 적절성, 응답내용의 비밀보장 준수여부, 조사원의 성실성, 휴식원칙 준수 여부 등등에 대해 정확한 기준을 갖고서 관리/감독함은 물론 단체별/개인별 조사수행을 객관적으로 평가를 해낸다. 이러한 평가를 통한 사후 포상제도를 둠으로써 보다 원칙을 지키려는 성실조사 유발을 꾀하는 것도 사소하지만 중요한 한 부분으로 인식되고 있다.

<표 4> 서울지역의 역대 출구조사 투표소 추출오차 및 실사오차 비교

	16대 총선 (3개 선거구 평균)	3회 지방선거	16대 대선
투표소 추출 오차	1.4	2.0	0.0
실사 오차	2.2	0.8	0.2

- 주) 1. 투표소 추출오차: | 추출투표소의 실제결과 - 전체 선거구의 실제결과 |
 2. 실사오차: | 추출투표소의 실제결과 - 출구조사 조사결과 |

3. 출구조사의 제약 조건

출구조사의 정확성만이 아니라 출구조사 수행 자체를 위협하는 제약조건이 있다. 다름 아닌 공직선거 및 선거부정방지법 167조에 규정되어 있는 300m 거리제한 규정이다. 이 제약 조항에 근거해 출구조사를 원천봉쇄하려는 선관위 관계자도 있으며, 투표소 접근 제지로 인해 응답자 선정의 체계성이 훼손될 수 있으며, 응답자와의 접촉자체의 제지로 조사가 중단되는 경우도 발생한다. 투표소를 충분히 많이 추출하지 못하는 경우에는 조사가 중지되는 투표소 한 두 곳이 전체 결과에 당연히 큰 영향을 미치게 된다. 전국적인 출구조사가 처음으로 도입된 지난 16대 총선의 경우 선거 관계자의 제지 사례율이 적게는 40%에서 많게는 57%까지 이르는 것으로 나타났으나(한국방송공사, 2000), 지난 16대 대선 때에는 그 제지가 줄어들기는 했지만 여전히 출구조사가 자유로운 환경에서 정상적으로 수행된 것은 아니다.

따라서 보다 정확하고 안정적인 조사결과를 얻기 위해서는 출구조사의 전면적 허용이

필수적이다.

<표 5> 16대 대선 문제발생 투표소 비율(미디어리서치 자체자료)

	조사 투표소 수	제지로 조사중지	문제발생 비율
수도권	77	8	10.4
충청권	31	1	3.2
호남권	25	3	12.0
TK권	21	5	23.8
PK권	31	2	6.5
강원/제주권	14	3	21.4

IV. 16대 대선 예측의 출구조사

1. 표본 추출

16대 대선에서 미디어리서치가 사용한 투표소 추출방법은 다단 층화 추출법이다. 먼저 1단계로 16개 시도를 하나의 층으로 간주하여 층화한 후, 16개 각 층 내에서 지난 3회 지방선거 자료를 중심으로 2~9개까지 층을 만드는 2단계 층화를 실시하였다. 여기서 중요한 연구성과가 하나 반영되는데, 바로 층화지표로 무엇을 사용하는가가 아주 중요한 부분이 된다. 미디어리서치가 층화지표로 삼은 것을 사정상 밝힐 수는 없지만, 일반적으로는 단순한 여당의 득표율 혹은 야당의 득표율, 여야간 지지도 차이(gap) 등등을 지표로 사용할 수 있을 것이다. 2단계 층화에서 또 하나 중요한 것은 층의 구성요소를 읍/면/동(邑/面/洞)으로 하느냐, 아니면 투표소를 기준으로 하느냐의 문제이다. 동(洞)을 기준으로 구성하면 동(洞)을 집락으로 추출하고, 다시 동(洞)내에서 최종투표소를 추출하는 방법이 될 것이며, 투표소를 구성요소로 하면 그 투표소 자체가 추출단위가 될 것이다. 미디어리서치는 투표소를 구성요소로 하여 층화하는 방식을 택하였다.

그 다음으로 2단계의 층내에서 표본 투표소 추출은 각 층 내에서 단순무작위로 추출하였고, 층별로 최소 2개 이상의 투표소가 추출되도록 하였다. 1개의 투표소만 추출할 경우 그 투표소가 선거 당일 문제가 발생할 경우 해당 층의 결과가 모두 없어지는 문제점에 대비하기 위해 최소 2개 이상의 투표소를 추출하였다. 한편, 층내에서 투표소를 추출할 때 층의 대표값에 가까운 투표소를 추출하는 대표구 방식 등을 적용할 수도 있다.

이런 과정을 거쳐 1차적으로 투표소 추출을 끝내고 실사 진행에 문제가 없을 지를 철저한 사전 답사를 통해 미리 점검하였다. 통계적으로는 문제가 없을지라도 거리제한 규정과 투표소 위치 환경 등 정확한 실사수행이 어려운 투표소가 발생시, 이를 동일 층내의 다른 투표소로 대체하는 과정을 거쳐 199개의 표본 투표소를 최종 확정하였다.

표본 투표소에서의 조사용답자 선정은 역시 체계적인 계통(systematic) 추출방식을 적용하였다. 즉, 성, 연령, 직업 등 일체의 응답자특성 변수를 고려하지 않고 오로지 체계적으로 매 k번째 응답자마다 조사를 하였다. 지난 대선에서는 그 k를 6으로 정했다. k의 간격은 목표 표본수와 그에 따른 통계량의 오차한계, 조사원수 등에 따라 다르게 정해지지만, k의 간격이 너무 넓거나 좁은 것은 바람직하지 못하다. 너무 넓은 경우에는 면접원의 집중력 저하로 간격을 놓칠 가능성이 높으며, 특히 투표율이 낮은 선거에서는 이러한 현상이 더 심해질 가능성이 높다. 또한 너무 좁을 경우에는 면접원이 너무 바빠 그것을 지키지 못해 체계적으로 응답자를 선정하지 못할 수 있다. 따라서 k의 간격을 정할 때는 통계적인 문제뿐만 아니라 면접원의 실사상황까지 고려해서 정해져야 한다.

<표 6> 16대 대선 미디어리서치(KBS) 투표소 추출 현황

시도명	투표구 수	총 수	추출투표소 수
서울특별시	2,214	9	34
부산광역시	902	5	15
대구광역시	578	2	9
인천광역시	573	2	10
광주광역시	323	3	9
대전광역시	311	3	10
울산광역시	267	2	6
경기도	2,358	8	33
강원도	738	3	9
충청북도	491	4	10
충청남도	727	4	11
전라북도	758	3	8
전라남도	997	4	8
경상북도	991	4	12
경상남도	994	3	10
제주도	249	2	5
계	13,471	61	199

2. 실사 진행

실사과정의 첫 번째 단계는 조사원 리쿠르팅과 조사원 교육이다. 지난 대선때에는 각 투표소별로 조장 1인과 조원 3~5명씩을 배정했는데, 출구조사에서 조장의 역할은 매우 중요하다. 조장이 역할을 잘못할 경우 조원전체가 잘못된 방향으로 조사를 할 수 있기 때문이다. 그리고 선관위 직원 제지 등의 문제가 발생할 경우에는 1차적으로 조장이 해결을 해야 하기 때문이다. 따라서 조장 선발은 출구조사 유경험 면접원, CLT 전문 면접원 등 유사 조사경험이 있는 사람을 중심으로 199명을 선발하였다. 조원의 경우도 출구조사 유경험 면접원을 우선적으로 선발하고 그 외의 인원은 자문 교수단의 추천, 대학 게시판 공고 등을 통해 모두 815명이 모집 동원되었다.

면접원 리쿠르팅이 끝난 뒤 교육은 2차례에 걸쳐 실시되었다. 첫 번째 교육은 출구조사의 중요성, 조 위치선정, 응답자 선정, 주의사항 등 이론과 원칙 중심으로 교육을 실시하였다. 출구조사 이론교육을 통해 내용을 전달하는 것도 중요하지만 면접원들이 아주 중요한 역할을 하고 있다는 자긍심을 고취시키는 것도 중요한 교육목적 중의 하나이다. 2차 교육은 사례위주의 교육 및 Role-playing 실시 등 실습중심으로 실시했다. 미디어리서치는 자체 평가자료와 참관인단 보고서의 자료를 활용하여 과거 출구조사에서 문제가 된 사례를 제시하고 각 사례마다 어떤 식으로 대처할 것인가를 교육시켰다. 그리고 조사원과 응답자의 역할을 교대로 시켜봄으로서 실사원칙을 숙지시켰다. Role-playing 실시를 통해 면접원들이 잘못 알고 있는 부분을 확인하고 정정시킬 수 있으므로 교육효과 확인 차원에서도 꼭 필요한 과정이라 할 수 있다.

자료 수집은 비밀보장을 위해 응답자가 직접 질문지에 기입하고 투입하는 투표방식(Ballot Method)을 적용하였는데, 면접방식보다 투표방식이 거절/무응답이 적다는 연구결과가 있으며(Bishop and Fisher, 1995), 우리의 몇 차례에 걸친 경험에서도 응답자의 비밀보장이 아주 중요한 요소임을 확인한 바 있어 응답자가 응답을 기입하는 동안에는 면접원이 비켜서서 한발 물러나 있는 등의 작은 부분까지 배려하였다.

투표방식을 사용할 경우에도 여전히 응답거절자의 문제는 남게 된다. 지난 대선의 거절자 분석결과 고연령·여성층에서 응답거절의 비율이 높게 나타났다(<표 7>).

<표 7> 16대 대선 출구조사에서의 성·연령별 거절률

	계	20대	30대	40대	50대	60세 이상
남자	15.8	8.4	10.4	13.6	22.4	26.7
여자	21.8	11.2	14.3	20.5	30.5	38.1
계	19.3	10.0	12.4	17.0	26.0	32.4

고연령·여성층과 같은 특정 계층의 표본 불포함률이 높을 경우에는 결과에 체계적인 편향이 발생할 수 있다. 따라서 출구조사의 조사거절자에 대해서도 성, 연령 등 기본적인

응답자 특성을 파악하는 것은 중요한 문제이다. 지난 대선에서 미디어리서치는 거절자일 경우 성은 관찰 가능하므로 연령만이라도 질문토록 하고, 연령마저 무응답일 경우에는 면접원이 주관적으로 파악하게 하였다. 이 때 면접원들의 주관적 연령 추정은 다소 오차가 있을 수 있겠지만 대략적으로 10세 간격 연령대만을 분석에 포함시키므로 거절자의 성·연령 응답자특성을 무시하는 것보다는 편향의 보정에 약간의 도움이 되었음을 분석을 통해 알 수 있었다(<표 8>).

3. 예측치 산출의 분석 과정

당일 출구조사 결과는 2시간 간격으로 집계가 되었다. 조사자료 전송방식은 ARS방식, 전화집계방식, PDA방식 등 여러 가지 방식이 있으며, 각 방식마다 장단점이 있다. 16대 대선에서는 전화로 결과를 불러받는 단순한 방식으로 자료를 집계하였다. 집계된 자료는 몇 가지 단계의 분석을 거쳐서 최종 예측치가 산출되었는데, 각 단계별 분석은 물론 사전에 모두 프로그램화되어 있어 큰 어려움은 없었다. 출구조사는 투표마감 시간까지 진행되었으나 17시까지 집계보고된 자료만으로 예측을 위한 분석에 들어갔다.

먼저 출구조사에 응하고 성·연령·후보지지 세 문항에 모두 응답을 밝힌 단순지지도 결과는 노무현 후보가 50.0%, 이회창 후보가 45.4%로 나타났다<표 8>.

이 단순집계 결과에 조사거절자/무응답층 보정을 하는 분석이 첫 번째 단계로 실시되었다. 성·연령만 파악되어 있는 조사거절자/무응답층의 사례수를 전체응답자에 포함시켜 성·연령층별 가중에 따라 전체응답자의 지지도를 다시 구했으므로, 조사거절 비율이 높은 계층에서 지지도가 상대적으로 높은 후보의 득표율이 약간 올라가는 경향을 보여준다. 보정 결과, <표 8>에서와 같이 노무현 후보 지지도가 47.0%로 다소 줄고, 이회창후보 지지도가 46.4%로 약간 올라가는 것으로 나타났다.

두 번째의 보정작업은 층별 크기를 고려하는 분석이다. 투표소 추출을 층의 크기에 비례하여 추출한 것이 아니므로, 정확한 층의 크기에 맞추어 가중치를 주는 분석을 실시하였는데, 그 결과 노무현 후보가 48.5%, 이회창 후보가 47.5%로 나타났다.

다음에는 지역별 예상투표율을 반영하여 전체득표율을 다시 계산하는 단계가 있었다. 일반적으로 전화조사인 경우에는 성·연령층별 예상투표율 가중치를 주어 분석하므로 투표율이 상대적으로 높은 고연령층에서 지지도가 높은 후보가 투표율 보정후 지지도가 약간 올라가는 경향이 있다. 그러나 출구조사의 경우는 투표자를 기준으로 조사하고 분석하기 때문에 성·연령별 투표율 가중치를 주는 과정은 필요가 없으며, 시/도 지역별 투표율만을 고려하면 된다. 지역별로 투표율이 차이가 있으므로, 지역별 예상투표율에 따른 예상 투표자수를 고려하여 전체득표율을 다시 계산하게 되면 투표율이 높은 지역에서 지지도가 높은 후보가 전체득표율이 약간 올라가는 경향을 보이게 된다. 이러한 투표율 보정분석

결과, 노무현 후보 48.7%, 이회창 후보 47.3%인 지지도를 얻었다.

마지막으로 부재자의 투표를 예측하여 반영하는 과정을 거쳤다. 부재자 수는 전체 유권자의 수에 비하면 그 비중은 그리 크지 않다. 그러나 부재자가 특정후보 지지성향이 강한 집단으로 구성되어 있을 경우에는 이들의 결과가 전체결과에 적지 않은 영향을 미칠 수 있다. 16대 대선의 부재자 867,476명은 전체 유권자(34,991,529명)의 2.5%에 해당되지만 부재자의 투표 결과는 1, 2위 후보 전체 표차 570,980표의 46.1%에 해당하는 263,354표로 나타나 이번 선거결과에 중요한 영향을 미쳤음을 알 수 있다. 따라서 부재자 부분을 제대로 반영하지 못할 경우에는 그 만큼 예측의 정확성은 떨어질 수밖에 없다. 이전 단계까지는 출구조사의 단순결과를 가지고 가중치를 부여하여 분석을 했지만, 부재자 부분은 별도로 부재자의 투표결과를 예측하여 반영해야 했다. 부재자 투표결과 예측방법 역시 전체 선거결과 예측의 중요한 한 부분이 되었다. 미디어리서치는 부재자의 구성 분포를 면밀히 분석하여, 과거 선거결과 및 사전 전화조사에서의 분석결과 등을 고려하여 부재자 결과를 거의 정확하게 추정하여 전체에 반영함으로써 최종 예측치인 노무현 후보 49.1%, 이회창 후보 46.8%의 지지도 값을 구할 수 있게 되었다.

<표 8> 16대 대선 출구조사 자료분석 단계별 결과 비교

보정분석 단계	노무현	이회창	차이
▶ 출구조사 단순집계	50.0	45.4	4.6
↓			
▶ 조사거절자/무응답자 보정	47.0	46.4	0.6
↓			
▶ 층별 크기에 대한 보정	48.5	47.5	1.0
↓			
▶ 투표를 보정	48.7	47.3	1.4
↓			
▶ 부재자 반영 ⇒ 최종 예측치	49.1	46.8	2.3
↓			
실제 결과	48.9	46.6	2.3

V. 결론

출구조사의 정확성을 안정적으로 확보하기 위해서는 몇 가지 검토할 사항들이 있다. 첫째로는 투표소 추출과 관련한 과제이다. 투표소를 추출하는 것은 출구조사의 첫 시작점이며, 조사결과의 승패에 가장 큰 영향을 미치는 요인이다. 단순랜덤, PPS방식을 제외한 다단계추출 등 나머지 추출방법은 과거 자료를 이용하여 주로 투표소를 추출하게 된다. 이때 과거의 많은 선거자료 중 어느 것을 사용할 것인가가 중요한 선택의 문제가 될 수 있다. 즉 최근의 선거를 사용할 것인가, 동일 성격의 선거를 사용할 것인가, 아니면 가능한 모든 선거를 통합해서 사용할 것인가의 문제이다. 이 부분에 대한 다양하고 지속적인 시뮬레이션 연구를 통해 각각의 선거에 있어서 가장 설명력이 높고, 보다 적합한 투표소 추출방법을 개발하는 것이 필요하다.

두 번째의 과제는 자료를 보정·분석하는 과정에 있다. 응답 거절자, 투표율, 부재자 등 단순 지지율 집계에서 제외되는 사람들을 어떤 식으로 분석하여 보정할 것인가에 대한 연구가 절대적으로 중요하다. 보수주의자 회피효과 등 특정 성향의 응답자가 체계적으로 응답거절이 나타난다는 연구가 일부 있지만, 그러한 효과를 정량적으로 측정하여 바르게 보정할 수 있도록 가중치 개발 등의 이론적 방법이 연구되어야 할 것이다.

효과적인 실사관리도 빼놓을 수 없는 과제이다. 실사관리는 전적으로 조사기관의 책임인 자질있는 조사원 확보와 내실있는 교육훈련, 원칙준수의 현장관리등과 관계가 있다. 그러나 기본적으로 선거가 없는 평상시에는 제한적인 인력자원으로만 운영될 수밖에 없는 조사기관들의 현실적 사정을 고려하면, 성공적인 선거예측을 위한 출구조사 준비를 위해서는 충분한 시간(time)적 여유와 예산(cost)적 여건이 감안되어야 함도 중요한 과제의 하나로 인식된다. 이러한 예산적 뒷받침은 안정적인 출구조사 투표소 수를 확보하는 결정적 요인이기도 하다. 홍내리 외(2001)에 의하면 2000년 16대 총선 출구조사의 사후 모의실험에서 투표소 수를 2배 정도로 늘릴 경우 예측 값들이 훨씬 안정적임을 확인할 수 있었다고 한다. 실제로도 미디어리서치가 16대총선에서 당선자 예측에 실패한 7곳 지역구 중 5개 추출이 3곳, 6개 추출이 2곳, 7개 추출이 2곳 이었으며, 9개를 추출한 3곳 지역구의 경우 당선자뿐만 아니라 득표율 예측도 정확했음을 알 수 있었다.

또한, 300m거리제한이라는 암적 조항은 재론할 여지없이 없어져야 할 걸림돌임에 틀림없다.

지금까지 우리나라에는 총 6회 정도 출구조사가 실시되었다. 국회의원 선거, 지방선거, 대통령선거, 재보궐 선거 등 각 선거별로 출구조사가 한 번 이상씩 실시되었는데 그 예측에 있어서 성공적인 부분도 있었고 다소 미흡한 점도 있었다. 그러나 출구조사에 대한 방법론이 제대로 정립되어 있지 못한 상황에서 현업실무자의 실험위주의 접근으로 설계되고 수행되었으며, 법제도적으로도 제한적인 환경에서 수행되어 온 것 또한 사실이다. 그런 상

황에서 출구조사가 나름대로 훌륭한 예측력을 보인 것은 매우 긍정적으로 평가받아야 한다.

그러나, 지금까지의 출구조사를 통한 예측성과가 지속적으로 유지 발전되고 한 단계 더 앞선 방법론으로 정립되기 위해서는 조사연구원, 조사기관들 뿐 아니라 학계 등 관련된 여러 분야에서의 끊임없는 연구가 뒤따를 때만이 가능하게 될 것이다.

참고 문헌

- 류제복. 2003. “출구조사의 역사와 개선방향.” 『조사연구』 4(1): 31-48.
- 홍내리·허명희. 2001. “제16대 국회의원 선거의 예측조사에 대한 사후적 검증.” 『조사연구』 2(1): 1-36.
- 한국방송공사. 2000. 『제16대 총선 당선자 예측조사 백서』. 한국방송공사.
- Bishop, G. F. and Fisher, B. S. 1995. “Secret Ballots and Self-Reports in an Exit-Poll Experiment.” *Public Opinion Quarterly* 59: 568-588.
- Marsh, Catherine., and Elinor Scarbrough. 1990. “Testing Nine Hypotheses about Quota Sampling.” *Journal of the Market Research Society* 32(4): 485-506.