

출구조사의 역사와 개선방향¹⁾

A history and the improvable direction of exit poll

류 제 복²⁾

Jea-Bok Ryu

본 논문에서는 선거예측 조사에 널리 사용되는 출구조사의 역사와 우리나라에서 실시된 출구조사의 문제점과 개선 사항들을 살펴보았다. 출구조사시 불충분한 준비로 발생할 오류들을 검토하였고 이들을 줄이기 위한 사항들을 추가적으로 고려하였다. 본 연구의 결과로부터 보다 정확한 선거예측을 기대한다.

We consider a history of exit poll which is generally used in the election forecasting survey. And we discuss some problems and improving items based on the results of exit poll executed in Korea. In addition, we consider the errors due to the insufficient preparation for exit poll and methods to reduce these errors. From this study, we expect the election forecasting will be more accurate.

I. 머리말

2002년도에는 몇 차례의 선거가 있었다. 6월 13일의 지방자치 단체장과 지방의회의원들을 선출하는 전국동시지방선거를 시작으로 12월 19일에는 제16대 대통령선거가 실시되었다. 선거예측을 위한 조사들이 많이 실시되고 다양한 방법들이 사용되었다. 특히 2000년 4월 13일에 실시된 16대 국회의원선거에서 처음으로 선을 보인 출구조사가 2002년도 선거예측조사에 더욱 빈번하게 사용되었다.

출구조사는 투표소에서 투표를 마치고 나온 유권자를 표본추출하여 어느 후보를 선택했는지를 조사하는 방법이다. 미국은 수정헌법 1조³⁾에 의해 보장된 제도이지만, 우리나라 통합선거법에서는 출구조사를 인정하면서도 투표소에서 30m 떨어진 곳에서

1) 이 논문은 2002~2003학년도 청주대학교 학술연구조성비(특별과제)에 의하여 연구되었음.

2) 청주대학교 생명·유전·통계학부 통계학전공(jbryu@chongju.ac.kr)

3) 미국 수정헌법 1조(1789년 헌법제정, 1791년 수정) : 미합중국의회는 종교를 수립하거나 종교의 자유로운 행사를 금지하거나 언론 또는 출판의 자유를 제한하거나 또는 평온하게 집회하고 고통의 구제를 위하여 정부에 청원하는 인민의 권리를 침범하는 법률을 제정할 수 없다

조사해야 한다고 규정함으로써 사실상 이를 금지하고 있다.

본 논문에서는 출구조사의 역사와 우리나라에서 실시된 출구조사의 결과를 분석하고 현행 출구조사의 문제점과 보다 정확한 출구조사가 실시되기 위해서 개선해야 할 사항들을 살펴보고자 한다.

II. 출구조사의 역사

출구조사를 초기에는 “선거날 조사”(CBS방송) 또는 “비밀조사”(NBC방송)라 하였는데 이 조사는 선거날 투표가 진행되는 장소에서 수행되는 표본조사이다. 출구조사의 장점은 투표가 진행되는 과정에서 자료가 수집되고 분석되기 때문에 다른 조사보다 훨씬 더 빠르고 정확하다는 것이다. 그러나 자료를 수집하고 전송하여, 통계처리 분석까지의 과정이 짧은 조사기간 내에 이루어져야 한다는 단점이 있다.

사전 전화조사가 출구조사와 다른 점은 사전 전화조사에 응답한 사람들 중에서 정작 투표를 할 때 지지하는 후보가 바뀔 수 있고 또한 투표를 할 것인지 안 할 것인지가 상황에 따라 변한다는 것이다. 더욱이 투표율이 낮거나 투표일 직전에 증대한 상황변화(2002년 12월의 16대 대통령선거에서 정몽준 대표의 노무현 후보지지 철회)가 생기면 이들을 충분히 반영하지 못한다.

1964년 메릴랜드 예비대통령선거에서 Louis Harris사의 Ruth Clark는 투표 당일 투표자들과 얘기를 나누고 싶어서 볼티모어로 갔는데, 그곳에서 투표자들을 만나는 일이 너무 힘들었다. 그래서 그녀는 투표가 진행되는 곳(공립학교)으로 찾아가서 선거관리자들의 허락을 받고 투표소를 나온 투표자들을 면담하였는데 이것이 출구조사의 시초가 되었다. Clark의 얘기를 들은 Harris사는 이 방법을 1964년 캘리포니아 공화당 예비선거에 사용하였다.

주요 언론매체에 의한 출구조사는 1967년 11월에 실시된 켄터키주의 주지사 선거에서 미국 CBS방송국에 의해 처음으로 사용되었다. 그러나 CBS는 1970년부터 1980년까지는 출구조사를 방송에서 단지 분석을 위해 사용하였고 1982년에야 비로소 출구조사를 예측에 사용하였다. 한편 WCBS-TV의 리포터 Earl Ubell은 1969년 뉴욕 시장선거에서 처음으로 출구조사에 대한 분석을 하였다. 출구조사가 사용된 지 35년이 지난 지금에는 출구조사가 주요 선거예측조사 방법으로 자리잡고 있다.

미국은 특성상 지역 간 시차가 커서 투표마감시간이 지역마다 다르다. 동부에서 조사된 결과를 바탕으로 발표된 조사결과가 서부의 선거에 영향을 미쳐 투표자들의 투표 성향을 바꾸게 하기도 하고, 투표하려는 사람들이 투표를 포기하게 하는 등 선거에 영향을 미치게 된다는 이유로 선거예측결과의 발표를 금지해야 된다는 주장이 제기되었다. 그러나 조사결과 그러한 주장이 사실이 아님이 밝혀졌다.

많은 의회 관계자들이 당선자가 선포되지 않은 투표마감 전에 선거결과를 예측하는 것을 싫어했지만, 방송사들은 1976년부터 1984년 사이에 출구조사에 의한 결과를 투표마감 전인 초저녁 뉴스에 방송하였다. 1984년까지는 법안은 아니지만 의회는 모든 주의 투표가 마감되기 전까지는 출구조사에 의한 선거예측을 금지하는 결의안을 통과시켰다. 출구조사 결과가 선거에 부정적 영향을 미치게 된다면 각 주는 출구조사를 반대하는 법안들을 만들기 시작하였다. 와이오밍주와 워싱턴주에서는 투표 장소로부터 300feet(약 90m) 이내에서 출구조사를 금하는 법안이 통과되었고, 이러한 추세가 플로리다주, 조지아주, 하와이주, 켄터키주 등의 여러 주로 확산되었다. 이에 대해 방송사들은 의회와 여러 주들의 반출구조사법에 대해서 출구조사에 의한 선거결과를 방송하였다. 마침내 1965년 Frank Stanton⁴⁾에 의해 처음으로 제안된 출구조사 법안의 장애물이 20년간의 논의 끝에 제거되었다.

방송사들은 각 주에서 제정한 반출구조사법에 대해 도전을 하게되는데 그 첫 번째 소송 대상은 워싱턴주이었다. 워싱턴주 연방법원은 출구조사가 투표소에서 평화와 질서, 그리고 예의를 파괴하는 지를 해당 주가 증명해야 한다고 하였다. 마침내 1988년 초에 방송사들은 법원으로부터 위헌 판결을 얻어내고 이어서 플로리다주, 조지아주, 켄터키주 등의 반출구조사법에 대한 도전에 성공하였다. 이를 계기로 캘리포니아주, 하와이주 등의 다른 주에서도 질서 있는 출구조사를 방해하지 않을 것에 동의하기에 이르렀다(Mitofsky, 1991).

Ⅲ. 한국에서의 출구조사

우리나라에서는 2000년 4월 13일에 실시된 16대 국회의원선거 예측에 처음으로 출구조사가 사용되었다. 그 후 2002년 6월의 전국동시지방선거, 8월의 국회의원보궐선거, 그리고 2002년 12월의 16대 대통령선거에 대한 예측조사에 전화조사와 함께 출구조사가 실시되었다. 특히 16대 대통령선거 예측조사에서는 방송 3사가 모두 조사전문기관과 함께 출구조사를 실시하였는데 출구조사결과는 16대 국회의원 선거에서보다 더 정확하였다.

2000년 4월 13일의 16대 국회의원선거에서 MBC가 한국갤럽과 KBS-SBS가 미디어리서치(MRI), 코리아리서치(KRC), 테일러넬슨소프레스(TNS), 그리고 한국리서치(HR)와 공동으로 처음 출구조사를 실시하여 선거예측을 하였다. 2002년 12월 19일에 실시

4) 방송사가 선거예측결과를 발표하는 것에 대한 공공의 불만을 알려주기 위해서 열린 1964-65년의 의회 청문회에서, 콜롬비아 방송사 사장인 Frank Stanton은 의회가 방송국이 예측결과를 방송하는 것을 간섭하지 말고 모든 주가 동일시간에 투표를 마감함으로써 문제를 해결하도록 하자는 제안. 당시에는 출구조사가 사용되지 않았지만 후에는 출구조사가 공격의 주요 대상이 되었다.

된 16대 대통령선거에는 MBC는 KRC, KBS는 MRI, 그리고 SBS는 TNS와 공동으로 출구조사를 실시하여 선거마감 직후인 6시에 3사가 동시에 출구조사결과를 발표하였다. 출구조사결과는 방송사간 다소의 차이가 있었으나 과거의 출구조사보다는 예측결과가 정확하였다. <표 1>은 출구조사를 처음 실시한 2000년 4월의 16대 국회의원선거와 최근의 2002년 12월의 16대 대통령선거에서의 출구조사 결과에 대한 비교표이다.

<표 1> 16대 국회의원선거와 대통령선거에서의 출구조사 현황

구 분	16대 국회의원선거	16대 대통령선거
조사대상	77~79개 선거구	투표소 : MBC(302), KBS(199), SBS(113)
조사시간	06시(또는 08시) ~15시(또는 16시)	06시~16시(또는 17시)
응답자 선정	5~7명당 1명씩 계통추출	6~12명당 1명씩 계통추출
조사방법	직접질문(MBC)과 투표형식(KBS-SBS)	투표형식
조사원	대부분 20대 여대생	KBS는 대학생과 기존의 면접원 등 여러 계층
조사원 배치	투표소 당 2~6명	투표소 당 4~6명
무응답률	20~30%	10~20%
무응답자 조치방법	교체하거나 성별, 연령만을 기입하여 표본에 포함시킴	성별, 연령 등을 기입하여 표본에 포함시킴
무응답 층과 출구조사의 이해 정도	40대 후반 이후의 연령으로 여성이 많음. 출구조사에 대한 이해가 부족하고 비밀보장을 우려	40대 후반 이후의 연령으로 여성이 많음. 출구조사에 대한 이해도는 많이 향상됨

(주) 무응답에는 응답거부가 포함되고, 앞으로도 같은 의미로 사용한다.

2000년 4월의 16대 국회의원선거 예측을 위해 실시된 출구조사는 전체 227개 선거구 중에서 MBC는 79개 선거구에서 그리고 KBS-SBS는 77개 선거구에서 실시되었다. 사전조사에서 경합지역으로 예측되거나 사전조사결과와 예상되는 결과와 차이가 큰 지역들을 출구조사 대상지역으로 선정하였다. 한 선거구에서 출구조사를 실시하는 투표소는 6개 투표소에서부터 21개까지 차이가 많았다. 투표소를 추출하는 데는 가가

운 시점의 역대선거에서 지역구 전체 결과와 각 투표소별 결과 차이가 가장 작은 투표소 순으로 추출하는 '역대선거오차의 최소제공법'이라 불리는 대표투표소 선출방법을 기준으로 하고 지역구내 특성, 층(strata) 설정법, 유권자수, 확률비례추출법 등의 다양한 방법들을 병행해서 투표소를 결정하였다. 조사 대상자들은 조사기관들 간에 약간의 차이는 있으나 대체로 5~7명당 1명씩을 계통추출하였다. 조사방법은 한국갤럽은 조사원이 직접질문해서 응답결과를 무선인터넷폰을 사용하여 본사 컴퓨터실로 전송하는 방법을 사용하였고 다른 4개 조사기관들은 응답자들에게 설문지를 주어 응답자들이 직접 기입해서 수거함이나 가방에 넣는 투표형식의 방법을 사용하였다. 투표소 당 조사원들은 1개 조당 한국갤럽이 2명, 다른 조사기관들은 4~6명이 배치되었다. 조사에 응답을 거절한 비율은 대략 20%~30%정도이고, 이중에서 몇 개 조사기관에서는 무응답자들의 성과 연령을 기입하거나 받아내기도 하였다. 무응답자들의 대부분이 40대 이후의 여성들이 많은 것으로 파악되었다. 한편 조사원들의 연령은 20대로 극히 일부를 제외하고는 여대생들이었으며 이들에 대한 사전교육을 2~6시간 정도 실시하였다. 출구조사에 대한 이해 정도는 낮은 편으로 특히 무응답계층에서는 출구조사에 대한 이해 부족이 무응답의 주요 원인이 되었고 응답에 대한 비밀보장을 우려하는 것도 하나의 원인이 되었다.

국회의원선거에서는 선거구별로 출구조사 대상지역을 선정하지만 대통령선거의 경우는 전국에서 대상지역을 선정한다. 2002년 12월에 실시된 16대 대통령선거 예측을 위한 출구조사를 16대 국회의원선거의 경우와 비교하여 차이가 나는 점들을 살펴본다.

조사시간은 출구조사가 처음으로 실시된 16대 국회의원 선거에서보다 16대 대통령선거의 경우가 길어져서 투표에 참여한 사람들이 가급적 표본에 반영되도록 하였다. 투표가 시작되는 06시부터 16시 또는 17시까지 실시되었고 MRI의 경우는 17시까지의 출구조사 결과를 사용하였고 향후 연구 분석을 위해서 실제조사는 선거마감시간인 18시까지 하였다. 응답자 선정은 16대 국회의원선거보다 추출간격이 길었으며, 조사방법은 이번의 대통령선거의 경우 3개 방송사 모두 실제 투표형식의 자기기입식 방법을 사용하여 응답자들의 신분을 최대한 보장하였다. 조사원은 국회의원선거에서는 대부분이 20대 여대생이 주축을 이루었다. 그러나 MRI는 류제복(2002년)이 제안한 바와 같이 남녀 대학생은 물론 기존의 면접원들을 조사원으로 활용해서 가급적 조사원의 분포를 다양하게 하였다. 출구조사에 대한 이해의 정도가 비교적 높아졌고, 무응답률도 낮아졌으며 무응답자들에 대한 처리 방법도 이번의 경우는 다른 응답자들로 교체하지 않고 성, 연령 등 파악할 수 있는 모든 자료들을 기재하여 표본에 포함시켰다.

IV. 출구조사에서의 문제점

마케팅조사, 여론조사, 실태조사 등의 각종 조사에 전화조사가 널리 사용되고 있다. 특히 선거예측조사에 전화조사가 많이 사용되고 있는데, 우리나라의 경우도 선거 몇 개월 전부터 전화조사를 통해 후보자나 정당 지지에 대한 조사를 실시하고 있다. 그러나 최근 들어서 전화조사의 응답률이 낮아지고 있으며 전화조사에 응답한 사람들 중에서 많은 이들이 실제 투표에 참여하지 않아 전화조사에 의한 예측결과의 정확도가 떨어지고 있다. 중앙선거관리위원회(<http://www.nec.go.kr>)의 자료 <표 2>에 의하면, 전국동시지방선거의 투표율이 가장 낮고 다음이 국회의원선거 그리고 대통령선거의 순이다. 과거보다는 현재에 가까울수록 투표율이 하락하고 있다. 92년에 실시된 14대 국회의원선거의 투표율은 71.9%이었으나 2000년도의 16대 국회의원선거에서는 투표율이 57.2%로 급격히 하락하였다. 한편 87년도에 실시된 13대 대통령선거에서는 89.2%의 비교적 높은 투표율을 기록하였으나 97년도의 15대 대통령선거에서는 80.7%이었고 16대의 경우는 15대보다 9.9%가 하락한 70.8%에 그쳤다. 이와 같이 각종 선거에서 투표율이 하락하고 있는 상황에서 실제 선거에 참여하지 않은 사람들을 대상으로 실시한 조사결과는 정확도가 떨어질 수밖에 없다. 투표율이 떨어지면 정확도를 유지하기 위해서 더 많은 표본이 요구된다. 따라서 실제 투표를 한 사람들을 대상으로 한 출구조사가 선호된다.

<표 2> 각종 선거에서의 투표율 변화

선거 종류	전국동시지방선거		국회의원선거(총선)			대통령선거(대선)		
	제1회 (95.6.27)	제2회 (98.6.4)	제14대 (92.3.24)	제15대 (96.4.11)	제16대 (00.4.13)	제14대 (92.12.18)	제15대 (97.12.18)	제16대 (02.12.19)
투표율	68.4%	52.7%	71.9%	63.9%	57.2%	81.9%	80.7%	70.8%

16대 국회의원선거에 대한 홍내리·허명희(2001)의 전화조사와 출구조사의 당선자 득표율 예측결과 비교를 보면, 전체적으로 출구조사가 전화조사에 비해 정확하다고 볼 수 있다. 그러나 이러한 비교는 조사대상 지역이 다르므로 정확한 비교가 곤란하다. 따라서 두 방법의 보다 정확한 비교를 위해서 출구조사가 실시된 지역에 대해 두 조사방법의 정확도를 비교해보고자 한다.

16대 국회의원선거에서 출구조사를 실시한 5개 조사기관 중 3개 조사기관의 조사방법별 단순집계치와 당선자의 실제 득표율에 대한 평균편차의 결과가 <표 3>에 있다. 조사기관별 예측치는 다를 수 있다.

비교대상은 출구조사가 실시된 선거구에 대해 출구조사 결과와 사전에 실시한 전화조사결과를 비교하였다. 각 조사기관은 투표일 전에 통산 2~3차례의 전화조사를 실

시하여 후보자들에 대한 득표율을 예측하였는데, 그 중에서 선거 당일에 가장 가까이 실시된 전화조사 결과를 비교대상으로 하였다. 대체로 최종 전화조사는 실제 투표일 1~3일 이전에 실시되었으므로 투표일에 가장 근사한 시점에서 조사된 것이다. 전화조사와 출구조사의 표준오차가 다르므로 단순한 편차의 비교는 곤란하지만 전반적인 경향을 보면 출구조사가 전화조사보다 편차가 훨씬 적음을 알 수 있다.

사전전화조사의 평균편차가 출구조사의 평균편차보다 큰 것은 출구조사 대상지역을 선정할 때 사전 전화조사의 예측치와 예상되는 결과의 차이가 큰 지역을 표본으로 선정한 것이 하나의 이유로 볼 수 있다.

<표 3> 16대 총선에서 조사방법별 단순집계치에 대한 평균편차
(단위 : %p)

조사기관	실제-전화조사	실제-출구조사
A(20개 지역)	11.89	3.06
B(18개 지역)	10.88	2.21
C(20개 지역)	11.33	2.74
전체(58개 지역)	11.38	2.68

<표 4>는 16대 대통령선거에 대한 단순예측치를 판별분석 등을 사용해서 조정한 예측치들의 절대편차로, 작을수록 예측의 정확성이 높게 된다. 선거결과 실제 득표율은 노무현 후보가 48.9%, 이회창 후보가 46.6%를 얻었다. 당선자의 실제 득표율과 선거를 앞두고 최종적으로 실시한 전화조사(통상적으로 1~2일 전에 실시한 자료임)와 당일의 출구조사에 의한 예측치들의 절대편차를 비교하였다. SBS-TNS는 당초 출구조사를 위한 모형을 설정하고 이를 적용할 예정이었으나 선거전날 정몽준 대표의 노무현후보 지지철회 발표로 당초 설정한 모형을 제대로 사용하지 못하고 선거 당일의 전화조사 결과에 의존하였다.

<표 4> 16대 대선에서 조사방법별 절대편차
(단위:%p)

방송사(조사기관)	실제-전화조사	실제-출구조사
갑	3.1	0.3
을	0.3	0.2
병	0.6	0.7

3개 방송사들의 예측치들은 허용오차가 다르므로 직접적인 비교가 곤란하다. 또한 조사결과로부터 얻은 단순집계치들을 바탕으로 방송사별로 서로 다른 조정방법을 적용하여 예측치들을 계산하였으므로 어떠한 조정방법을 사용하였는가에 대한 검토가 필요하다.

<표 3>과 <표 4>로부터 출구조사의 예측 정확도가 전화조사보다 높음을 알 수 있다. 따라서 향후 선거 예측조사에 출구조사의 사용은 계속 증대할 것이다.

우리나라에서 실시된 출구조사의 결과를 분석해 본 결과 다음과 같은 문제점들이 제기되었고 이에 대한 개선방법을 살펴보았다.

① 법적문제 : 출구조사는 투표를 하고 나오는 사람들을 대상으로 조사를 하는 것인데, 우리나라 선거법 제 167조 2항에 의하면, 선거일에 투표소로부터 300미터 이내에서는 조사가 금지되어 있으므로 실질적인 출구조사가 불가능하게 된다.⁵⁾ 따라서 출구조사가 널리 사용되고 있는 나라에서와 같이 우리나라에서도 투표소로부터의 거리 제한을 없애거나 완화해야 한다.

② 조사원의 선정과 훈련 : 16대 총선과 대선의 출구조사에서 무응답은 40대 후반 이후의 연령층에서 높았고, 특히 여성이 더 많은 비중을 차지했다. 대도시보다는 중소도시의 경우가 응답거절율이 높았다. 대부분의 조사기관들은 20대 여대생을 조사원으로 채용하였는데, 이는 재고의 여지가 있다. 가령 5명을 1개조로 편성하는 경우 연령대별로 조사원을 구성해서 응답자들의 연령에 가까운 사람이 조사에 임하는 것이 바람직하다. 출구조사원들에 대한 교육은 2~6시간 정도로 실시하였는데, 응답자 선정에 사용된 계통추출법의 중요성, 무응답자 처리방법, 그리고 조사에 대한 충분한 이해를 시키는 위해서는 과거보다 더 많은 교육시간을 할애해야 한다.

③ 투표소 수와 투표소 선정방법 : 몇 개의 표본 투표소를 선정하는가의 문제는 추정의 정도와 관계가 있다. 표본투표소의 선정은 대부분이 역대선거오차가 최소가 되는 대표투표소를 추출하는 비확률적 방법을 사용하였다(홍내리와 허명희(2001)는 모의 실험을 통해서 투표소 수와 투표소 선정이 부적절하였음을 보여주고 있다). 물론 일부에서는 확률비례추출법(probability proportional sampling)이나 다단계층화집락추출(multi-stage stratified cluster sampling)과 같은 확률추출법을 사용하기도 하였다. 통상적으로 표본투표소의 추출은 과거 선거에서 투표에 영향을 미치는 변수들을 사용해서 층화한 후(예를 들어, 아파트 지역과 아닌 지역 또는 상가 지역과 아닌 지역 등으로) 층내에서 확률표본(random sample)을 사용하는 것이 바람직하다. 이때 투표소 별로 유권자 수가 동일하지 않으므로 유권자 수를 바탕으로 한 확률비례추출법을 사용하는 것이 바람직하다. 그래야 추정의 정확성을 평가할 수 있다.

5) 2000년 2월 16일 선거법이 개정되기 전에는 투표소로부터 500미터 이내에서는 조사가 금지되었다.

④ 표본으로 선정된 투표소 내에서 응답자들의 선정 : 표본 응답자들의 선정은 계통추출법을 사용하였고 계통추출을 적용하기 위한 추출 간격도 거의 동일하였다. 그러나 추출 간격의 결정은 각 투표소마다 요구되는 표본 크기와 예상되는 투표자수에 따라 달리 결정되어야 한다. 왜냐하면 투표소마다 유권자 수가 다르고 예상 투표율이 다르므로 이를 감안하고 추정오차의 정도를 함께 고려해야 한다.

⑤ 무응답자들에 대한 조치 : 16대 총선에서는 무응답자들이 생기면 이들을 다른 사람으로 교체하거나 아예 제외하는 방식을 사용하였다. 그러나 16대 대선에서는 무응답자들의 성별, 연령 등을 기재하고 표본에 포함시켰다. 원천적으로 무응답이 발생하지 않도록 하는 것이 중요하나 무응답이 생기면 이로 인해 추정의 편향이 생기는 것을 방지하기 위해서 다양한 조치를 취해야 한다. 16대 대선에는 항목 무응답이 다수 발생한 것으로 보이므로 이에 대한 대체방법(imputation)의 연구가 필요하다. 무응답은 대부분이 조사원이 응답자를 접촉하는 순간에 발생하나 설문지 길이에 비례하기도 한다.

⑥ 조사방법의 문제 : 16대 총선의 경우는 직접질문 방식과 자기기입방식이 함께 사용되었는데 16대 대선에서는 모두 자기기입방식에 실제 투표하는 것과 같은 형식을 사용하였다. Bishop과 Fisher(1995)의 연구에 의하면 출구조사에서 대면조사보다 비밀투표형식의 조사가 무응답이나 응답 거절 비율이 낮음을 알 수 있다. 출구조사에서는 자기기입식 비밀투표형식이 바람직하다 하겠다.

국회의원선거와 같이 지역 단위 선거예측에 사전전화조사를 통해 출구조사 대상지역을 선정하는 것은 문제가 있을 수 있다. 왜냐하면 최근 들어서 투표율이 낮아짐(<표 2>)으로 전화조사에 응답한 사람들(유보층도 포함) 중에서 실제 투표에 참여하지 않는 경우가 많이 발생하므로 사전에 철저한 분석이 필요하다.

V. 개선 방법

우리나라에서 실시된 출구조사의 문제점과 이를 개선하기 위한 방법에 대해 IV장에서 살펴보았다. 류제복(2000)과 홍내리·허명회(2001)도 출구조사의 정확성과 효율성을 높이기 위한 여러 방안들을 제시하였으며, 특히 이창현(2002)이 제안한 예측조사에서 선거자문·감리체제의 도입은 바람직하고 본다. 본 절에서는 추가적인 사항과 재고해 볼 사항들을 간추려 보았다.

1. 조사방법

① 표본추출 : 투표소(선거구)의 유권자 수에 비례하는 확률비례추출이 바람직하다.

그리고 투표소(선거구)의 영역을 가장 최신의 것으로 수정한다. 응답자의 익명성을 보장하기 위해 응답자를 선정할 때도 확률추출을 적용한다.

② 설문지 : 설문지의 길이가 짧을수록 응답률은 높게 된다. 가장 짧은 것은 누구에게 투표하였는지를 묻는 것이다. 좀 긴 설문지에는 어떤 주제에 대한 응답자의 태도, 그 밖의 의견, 그리고 인구학적 특성들을 포함시켜서 가중치 적용이나 투표소별 특성 파악 등의 향후 연구에 사용한다. 서술형 질문은 피하고 자기기입식 방식을 사용한다.

③ 조사 : 경험이 풍부한 조사원이 조사를 담당하는 것이 바람직하고, 조사원들에게 방송사의 배지나, 옷 등을 입게하여 신뢰할 수 있도록 한다. 무응답자들에 대해서는 성별, 출신지역, 또는 연령 등과 같은 가능한 모든 것을 기입한다. 이는 무응답 층에 의한 편향을 조정하는 데 사용한다.

④ 자료평가 : 전체적으로 출구조사의 표본추출오차는 동일한 표본투표소의 실제 투표결과를 기초로 한 추정치의 표본추출오차에 비해 크게 된다. 이는 투표한 사람들로 부터 표본을 추출함으로써 생기는 추출오차와 무응답들에 의한 편향에 기인한다. 조사의 결과보다는 어떻게, 왜 그런 결과가 나왔는지를 파악하기 위해서 조사기관과 전문가들과의 공동연구를 통해서 원자료에 대한 다양한 분석이 필요하다.

⑤ 조사시간 : 16대 대통령선거에서의 출구조사는 투표가 시작된 06시부터 16시~17시(투표마감은 18시)까지 실시되어 초기의 출구조사 시간보다 길어졌다. Busch와 Lieske(1985)들의 연구에 의하면 그룹별로 투표시간대에 따라 투표율에 유의적 차가 있음을 알 수 있는데, 우리나라의 경우도 MRI의 자료에 의하면 시간대별로 연령간 투표율에 차이가 있음을 알 수 있다. 특히, 16대 대통령선거의 경우는 연령대별로 지지하는 후보의 차이가 있었고, 투표시간도 연령이 많은 층은 오전에, 그리고 연령이 적은 층은 오후에 투표를 하는 경향이 두드러졌다. 16시에 출구조사를 마감한 KRC와 TNS는 17시까지 조사한 MRI에 비해 노무현 후보의 득표를 낮게 예측하였는데 이는 연령별 투표 시간대의 차이로 인한 것이 한 원인이라 볼 수 있다. 따라서 가급적 출구조사의 시간을 최대한 길게 하고 시간대별로 특정그룹의 투표형태를 분석하여 가중치 조정에 반영해야 한다.

2. 무응답 처리

출구조사에서의 무응답률은 16대 총선에서는 20~30% 그리고 16대 대선에서는 10~20%정도로 집계되었다. 무응답 집단이 응답집단과 차이가 나는 경우 각별한 주의가 필요하다. 조사기관에서는 표본으로 선정된 사람이 응답을 거부하거나 기피할 경우 표본을 교체하거나 이들을 제외하고 조사를 진행하였다. 무응답자를 교체하지 않을 경우에 무응답 집단이 응답집단과 다른 특성을 갖고 있다면 이러한 특성이 반영되

지 않아서 실제 투표율 계산에 편향이 생긴다. 또한 무응답자들을 교체하는 경우에도 이들 집단과 동일한 집단으로 교체하지 못하게 되면 역시 무응답자들의 특성을 반영하지 못하고 나아가 전혀 다른 특성을 갖고 있는 교체 집단의 특성이 그대로 반영되어 편향이 더 커질 수 있다(류제복, 2000).

1) 무응답집단을 교체하지 않은 경우

조사 대상 지역에서 특정 후보에 대한 모집단 지지율을 응답집단과 무응답집단으로 나누어 보면 다음과 같이 된다.

$$p = \lambda p_n + (1 - \lambda) p_r \quad (1)$$

여기서, λ 는 무응답집단의 비율이고, p_n 과 p_r 은 각각 무응답집단과 응답집단의 특정후보에 대한 지지율이다. 조사오차가 없는 경우를 가정하고, 무응답자들을 교체하지 않고 응답자들만을 대상으로 한 특정 후보의 지지율 p 에 대한 불편추정량을 \hat{p}_r 라 할 때 편향은 다음과 같이 된다.

$$\begin{aligned} Bias(\hat{p}_r) &= E(\hat{p}_r) - p \\ &= p_r - \lambda p_n - (1 - \lambda) p_r \\ &= \lambda(p_r - p_n) \end{aligned}$$

(2)

무응답률 λ 와 응답집단과 무응답집단의 특정후보에 대한 지지율의 차에 따라 편향이 발생한다. λ 와 비율 차($p_r - p_n$)가 클수록 편향이 증가하는 것을 쉽게 알 수 있다. 즉, 무응답 집단과 응답집단의 특성에 차이가 많이 날수록 응답자들만을 대상으로 예측한 추정치에 편향이 크게 된다. 따라서 무응답집단을 배제한 채 조사결과를 예측하여서는 안되며, 무응답 비율을 최대한 낮추도록 노력해야 한다.

2) 무응답집단을 잘못 교체하는 경우

무응답자들을 다른 표본으로 교체하여 투표율을 추정하는 경우, 1)절에서와 같이 λ 를 무응답집단의 비율이라 하고 무응답집단 대신에 교체한 표본에 의한 추정치를 \hat{p}_s 라 두면 응답자들과 교체응답자들을 기반으로 한 p 의 불편추정량 \hat{p}_{rs} 의 기대값은 다음과 같이 된다.

$$E(\hat{p}_{rs}) = \lambda p_s + (1 - \lambda) p_r$$

(3)

그러면 조사오차가 없다고 할 때, \hat{p}_{rs} 의 편향은 식(1)과 식(3)으로부터,

$$\begin{aligned} \text{Bias}(\hat{p}_{rs}) &= E(\hat{p}_{rs}) - p \\ &= \lambda(p_s - p_n) \end{aligned} \quad (4)$$

이 된다.

1)절에서와 마찬가지로 무응답집단을 다른 표본으로 교체할 때 무응답집단과 특성 차이가 큰 집단으로 교체하는 경우 편향은 증가하게 된다. 따라서 무응답집단을 교체할 때 무응답집단에 대한 면밀한 분석이 요구된다. 이를 위해서는 무응답집단에 대해 가능한 많은 정보를 수집하고 사전 조사나 과거의 조사를 바탕으로 유권자들에 대한 성별, 연령별 등의 다양한 특성들을 파악해서 무응답집단의 교체시 반영해야 한다.

아울러 응답자들이 올바른 응답을 하지 않는 경우도 고려해야 한다. 앞의 두 경우에서 보았듯이 응답률이 높을수록 추정의 편향은 감소하게 된다. 그러나 이를 위해서 무리하게 응답률을 높이려 응답자들에게 응답을 강요하는 경우 응답자들이 거짓 응답을 하게 된다. 그러면 추정시 문제가 더욱 복잡해진다. 거짓응답의 문제를 다루고자 할 때는 류제복 외(1993)의 책자를 참조하면 된다.

3. 가중치조정

선거 예측조사에서 얻어진 원자료(raw data)에 대해 통상적으로 다음 3단계를 거쳐 가중치를 조정해 준다(Levy, 1983).

1단계 : 투표소 내에서의 가중치로 응답자들이 선정될 확률과 투표소 내에서 무응답 비율을 감안한 가중치, 즉, $w_{i1} = \frac{N_i}{n_i}$. 여기서 N_i 와 n_i 는 각각 i 번째 투표소의 총 투표자수와 출구조사에 응답한 사람 수이다.

2단계 : 투표소가 선정될 확률을 감안한 가중치로, 표본 투표소의 추출확률은 투표소의 유권자수 등에 비례하는 불균등추출확률을 갖는다.

3단계 : 전체 표본이 집계된 투표소 수준의 통계표를 예상되는 투표결과에 따라 조정한다.

3단계의 가중치 조정을 반영한 다음과 같은 비추정치를 생각할 수 있다.

$$E = \left(\frac{\sum x_i}{\sum y_i} \right) Y$$

여기서,

E = 특정 후보에 대한 추정 최종 투표자수

x_i = i 번째 투표소에서 특정 후보에 대한 가중 투표자수

y_i = i 번째 투표소의 과거 선거에서 비교 가능한 후보에 대한 기지의 가중

투표자수(출구조사 전에 사전조사를 한 경우에는 사전조사 자료를 활용한다)

Y = 과거 선거에서 비교 가능한 후보에 대한 투표자수

위의 비추정치는 한 시점에서 특정 후보에 대한 투표결과의 추정치를 제공해주므로 이러한 과정은 모든 후보자들에 대해서 이루어져야 한다. 한편 특정한 성별이나 연령대의 사람들에 무응답이 많으면 성별, 연령 등을 고려한 가중치를 사용할 수 있다. 조성경(2002)은 전화조사, 모바일조사, 그리고 인터넷조사에서 가중치효과를 비교해 보았는데 출구조사에도 이러한 방법의 검토가 필요하다.

VI. 맺는말

본 연구에서는 출구조사가 오늘에 이르기까지의 역사를 간략히 살펴보았다. 우리나라에서는 출구조사가 2000년 4월에 실시된 16대 국회의원선거 예측에 처음으로 사용된 이후 최근의 2002년 16대 대통령선거 예측에 이르기까지 자주 사용되고 있다. 16대 국회의원선거와 대통령선거 예측에 사용된 출구조사의 결과를 바탕으로 출구조사의 전반적인 상황을 분석하고 제반 문제점들과 개선할 사항들을 검토하였다. 그리고 출구조사를 통해 보다 정확한 선거예측을 위해서 출구조사시 추가로 개선할 사항과 출구조사의 불충분한 준비로 인한 영향력 등을 알아보았다.

참고문헌

- 류제복. 1993. “대체 확률화응답기법.” 《응용통계연구》 6(2): 311-318.
- 류제복. 2000. “선거예측조사의 신뢰성 증진방안 - 16대 총선을 중심으로.” 《조사연구》 1(2): 15-34.
- 류제복. 2002. “출구조사의 문제점과 개선방향.” 한국조사연구학회 2002년 춘계워크샵 발표논문집: 79-93.
- 류제복·홍기학·이기성. 1993. 《확률화응답모형》. 자유아카데미.
- 미디어리서치. 2002.12. “미디어리서치-KBS 16대 대선 예측결과.”
- 박용치. 2000. “선거예측에서 편의의 감소: 거짓 응답을 중심으로.” 《조사연구》 1(2): 35-57.
- 이창현. 2002. “대통령 선거시기 여론조사의 기능과 역할에 관한 연구-KBS 방송사 예측조사 자문·감리 시스템을 중심으로.” 한국조사연구학회 2002년도 추계학술대회 발표논문집: 31-53.

- 조성겸. 2002. “새로운 사회여론조사 방법-모바일 조사방법의 전망.” 한국조사연구학회
2002년도 추계학술대회 발표논문집: 91-108.
- 중앙선거관리위원회. <http://www.nec.go.kr>
- 중앙일보. 2002년 12월 20일
- 홍내리 · 허명희. 2001. “제 16대 국회의원 선거의 예측조사에 대한 사후적 검증.” 《조
사연구》 2(1): 1-35.
- Bishop, G. F. and Fisher, B. S. 1995. “Secrete Ballots” and Self-Reports in an
Exit-Poll Experiment.” *Public Opinion Quarterly* 59: 568-588.
- Busch, R. J. and Lieske, J. A. 1985. “Does Time of Voting Affected Exit Poll
Results?.” *Public Opinion Quarterly* 49: 94-104.
- Levy, M. R. 1983. “The Methodology and Performance of Election Day Polls.”
Public Opinion Quarterly 47: 54-567.
- Mitofsky, W. J. 1991. “A short history of exit poll.”, in *Polling and Presidential
Election Coverage*, ed. Paul J. Lavrakas and Jack K. Holley. 83-99,
Newbery Park, CA;Sage.
- Sudam, S. 1986. “Do Exit Polls Influence Voting Behavior?.” *Public Opinion
Quarterly* 50: 331-339.
- 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.