

연구논문

대안적 사회여론조사 방법 :  
모바일 조사방법의 가능성 검토\*

A New Alternative Method for Social Survey: Possibility of Using  
Mobile Phone Survey Method

조성겸\*\* · 강남준\*\*\*

Sung Kyum Cho · Namjun Kang

모바일 조사방법은 유선전화조사와는 달리 응답대상자에 대한 접촉가능성이 높기 때문에 새로운 조사방법으로 활용되기 시작하였다. 그러나 모바일 조사방법의 표본 대표성을 확보할 수 있는 방법론이 아직 정립되지 않았다. 이 연구는 성향지수를 이용한 가중치를 사용하는 것이 모바일 조사의 표본편향을 제거해 줄 수 있는지 검토해 보았다. 성, 연령, 학력 등의 인구학적 속성과 기타 태도변인을 이용하여 성향지수를 산출하여 가중치로 사용하였다. 그 결과 선거예측조사와 과학 여론조사 등에서 표본의 편향이 감소되는 것으로 나타났다. 그러나 이러한 본 연구의 결과가 일반성을 어느 정도 가지는가는 좀더 검토되어야 할 것이다.

Telephone surveys miss, among other people, those who live in homes without telephones, people who are away from home at the time of interview and people who refuse to be interviewed. Recently, mobile phone survey has emerged as 'A replacement technology' to the old telephone survey method. Mobile survey enables us to do many things we could not do or could not afford to do before, and greatly enhance the efficiency of the opinion surveys. Very specifically, the mobile survey enables us to control respondent's accessibility, interviewer bias and to do incredibly fast and at a affordable costs. The authors analyze the results of mobile-phone local election polls and ELSI bio-technology attitude survey. The authors describe their results, the methods they used, including the use of demographic and propensity weighting to correct for substantial biases in the raw, unweighted data. The results show that mobile survey can predict the election outcomes with approximately equal accuracy to that of the telephone poll after weighting. This paper also cautions readers not to assume that mobile survey can be used with equal success in other elections and emphasizes the need for continuing research to improve mobile survey methods in the future.

\* 이 논문은 과학기술부 21세기 프론티어연구개발사업 중 인간유전체기능연구 사업의 'ELSI 연구'(FG-4-19) 지원에 의해 수행되었습니다.

\*\* 충남대학교 언론정보학과. skcho@cnu.ac.kr

\*\*\* 한양대학교 신문방송학과. njkang@hanyang.ac.kr

## I. 서론

### 가. 문제의 제기

최근 모바일 전화를 이용한 조사방법이 새로운 사회조사 방법으로 활용되기 시작하였다. 모바일 조사가 관심을 끄는 것은 유선전화와는 달리 응답대상자에 대한 접촉가능성이 높기 때문이다. 유선전화의 경우 재택자만을 접촉할 수 있기 때문에 재택률 차이로 인한 표본 편파가 발생한다. (조성겸 1997 ; 홍내리 허명희 2001). 그리고 통신매체로서의 비중이 유선에서 무선으로 옮겨지는 사회적 추세도 이러한 사회조사 방법의 변화를 촉진할 것으로 보인다. 사회조사 방법이 방문면접 또는 우편조사 방법에서 전화로 바뀌었는데, 이제는 다시 전화에서 새로운 테크놀로지를 사용하는 모바일 조사로 바뀌고 있는 것이다(Dillman 1998). 새로운 여론조사 기법의 등장과 쇠퇴 과정을 보면 시대적 변화에 적응하지 못하고 과거의 영화에만 집착했던 수많은 기업들이 도산하는 과정과 비슷한 것 같다. 첫 번째로는 1936년 루스벨트(Roosevelt)와 랜든(Landon)의 선거결과를 잘못 예측했던 'Literary Digest'가 1916년부터 대통령 선거 여론조사에서 써왔던 우편조사기법(mail survey)이 그 예이다. Literary Digest는 우편조사 방법을 사용해 대통령 선거결과를 정확히 맞추었고, 특히 1932년 선거에서는 아주 근소한 차이로 루스벨트의 승리를 예측했다. 그러나 1936년 선거에서는 우편조사 표본의 잘못된 표집 틀(sampling frame) 때문에 예측에 실패하게 되고, 이때부터 여론조사에서 우편조사 방법 대신에 호별방문에 의한 인터뷰 기법(door-to-door canvassing)을 주로 사용하게 되었다.

이어 1948년 미 대선이 또 한 번 여론조사 기법 변화의 고비가 되었다. 모든 여론조사회사들이 트루만(Truman)의 마지막 역전 상승세를 감지하지 못하고 듀이(Dewey)의 낙승을 예측했었다. 그 결과 의회 청문회까지 열려 결국 확률표본이론에 의한 표집기법이 자리 잡게 되었으며 응답자를 면접자가 선택했던 호별방문 인터뷰기법은 여론조사 시장에서 점차 사라지게 되었다. 그 후 가정전화의 보급이 보편화되면서 전화번호부를 표집 틀로 사용한 확률적 표본조사 방법이 신속성과 저렴한 비용 때문에

주된 조사방법으로 사용되기 시작했고 그 결과 전화조사방법이 현재 여론조사의 중심 조사기법으로 자리 잡았다.

그러나 모든 조사가 동일한 방법에 의해 수행되면 특정 조사방법에 따른 편향(bias)이 발생할 경우 모두 같은 굴레를 뒤집어쓰게 된다. 그 동안 사용되어 왔던 여론조사 조사기법이 그랬듯이 사회가 변천하면서 유선전화에 의한 여론조사 방법은 최근 여러 가지 문제점을 드러내기 시작했다. 요일 / 시간대 별 재택률 차이에 따른 표본편포 문제, 낮은 응답률 때문에 발생하는 표본의 대표성 문제, 개인의 투표성향을 조사하는 데 있어서 개인전화가 아닌 가구전화(household phone)의 사용 문제 등 체계적 편차요인이 존재한다고 밝혀졌다(홍내리, 허명희, 2001 ; Taylor, 2000).

그래서 여론조사 결과를 사용한 과학적 정밀 보도론의 창시자인 마이어(Philip Meyer)는 1990년대에 들어서면서부터 여론조사의 정확도 — 여기서 는 예측 득표율과 실제 득표율과의 차이 — 가 점점 더 커졌으며,<sup>1)</sup> 기존 전화여론조사 방법만을 고집하는 여론조사회사들은 2000년 미국대선 예측조사에서 고전할 것이라고 예측하였다. 반면, 대안적 조사기법인 인터넷 조사기법을 사용하는 Harris Interactive나 Knowledge Network의 여론조사 결과가 더 정확할 것이라고 전망하였다(USA Today, 10/31, 2000). 이러한 Meyer의 추정은 그대로 들어맞아 2000년 선거결과 최종 예측조사에서 유선전화 조사방법에 의존한 Gallup은 고어(Gore)의 예상 득표율을 실제 득표율보다 2.4% 낮게 보았다. Gallup의 라이벌인 Harris는 인터넷 조사를 기존 전화조사와 병행해 1.4%의 오차를 보였는데, 더 중요한 점은 Gallup이 고어(Gore)와 부시(Bush)의 득표율 차이가 2%(46% 대 48%)라고 최종 결론을 내린 반면, Harris는 0%(47% 대 47%)로 추정해 실제 차이(0.5%)에 아주 근접했다는 사실이다. 따라서 2001년 2월 미 하원에서 열

1) Gallup 조사의 경우 대통령 선거에서 승리한 후보자의 예측 득표율과 실제 득표율과의 차이 평균은 1936~1948, 4.3%; 1952~1968, 2.0%; 1972~1988, 1.6%로 점점 낮아지더니 1992~1996에는 3.3%로 다시 높아졌다. 2000년 대선에서 전국적 여론조사 결과를 발표한 19개 조사의 예측률 차이는 평균 1.7%로 다시 줄어들었는데, 이러한 감소효과는 당시 후보간 접전이 예상되어 표본크기를 전에 비해 크게 늘린 결과 발생한 것이다.

#### 4 조사연구

렸던 2000년 미국 대선 여론조사 문제점에 대한 청문회에서 언론매체는 여론조사 결과를 보도할 때 한 가지 방법으로 조사한 결과에만 의존하지 말고 다른 방법에 의해 조사된 결과도 함께 고려해 판단하는 것이 좋다고 결론을 내렸다.<sup>2)</sup>

2002년 11월 5일 실시된 미국 중간선거에 대한 여론조사 결과도 신통치 않았다. 2004년 대선을 목표로 조사기법 등을 재정비해 왔던 조사회사들은 2002년 선거예측 결과에 충격이 컸다. 그리고 2004년 대선 조사를 제대로 할지 의문이 간다는 비판이 제기되었다(New York Times, 11/6, 2002; Wall Street Journal, 11/7, 2002). 11월 5, 6, 7일 계속 연재된 New York Times의 분석기사에서 여론조사 예측의 실패는 기존 유선전화 여론조사가 가지고 있는 문제점, 즉 자동응답기에 의한 선택적 통화, 발신자 번호 추적 시스템의 보편화에 따른 미확인 번호에 대한 수신거부, 40대 이하 젊은층의 휴대전화 사용률 급증 등이 전체적으로 낮은 응답률을 가져와 여론조사 예측이 빗나가게 되었다고 하였다. 따라서 인터넷, 휴대전화 등을 사용하는 새로운 여론조사 기법의 개발이 중요하다고 주장했다. 이것은 2000년 대선 직후 여론조사에 대해 미 의회 청문회에서 내렸던 주류 전화 여론조사방법에 병행하여 대안적 조사방법을 사용해야 한다는 권고와 유사하다고 볼 수 있다. 이러한 새로운 대안적 여론조사방법 중 우리나라 상황을 고려할 때 우선적으로 고려할 수 있는 것이 휴대전화를 사용하는 모바일 조사이다.

그렇지만 이러한 모바일 조사는 유선전화의 경우 그랬던 것처럼 새로운 사회조사 방법으로 자리 잡을 수 있을지의 여부가 아직 명확치 않다. 전화조사가 1970년대에 미국에서 사회조사 방법으로 자리 잡을 수 있었던 것에는 전화를 이용한 조사가 저렴하고 신속했을 뿐만 아니라, 전화번호부를 이용함으로써 확률표집이 가능할 것으로 믿어졌기 때문이었다(Frankel and Frankel, 1987). 그러나 모바일 조사는 표본의 대표성을 확보할 수 있는 방법론이 아직 정립되지 않았다.

그 동안 모바일 조사방법에 대해서는 학술적 관심이 많지 않았다. 그래서 표본 추출방법은 표집 편의상 기존의 할당표집 절차를 그대로 이용하

2) <http://energycommerce.house.gov/107/hearings/02142001hearing216>

고 있지만, 과연 이러한 방법이 모바일 조사 상황에서 어느 정도 타당한지는 아직 검토되지 않았다. 간혹 모바일 조사결과를 전화나 다른 조사방법에 의한 조사결과와 비교하면서 방법에 따른 차이가 크지 않다는 점이 제시된 경우가 있지만,<sup>3)</sup> 이러한 결과가 모바일 조사의 타당성을 보여 주지는 못한다. 다른 조사방법 자체도 편향을 가지고 있기 때문이다. 이처럼 모바일의 표집에 대한 관심과 논의가 부족하기 때문에 아직 모바일 조사가 사회조사 방법으로써 어느 정도 타당성을 가지는지에 대해서 알 수 없을 뿐만 아니라 방법적인 측면에 대한 개선노력도 이루어지기 힘든 실정이다.

외국의 경우에도 일부 유럽국가를 제외하고는 모바일에 대한 연구 관심이 높지 않다. 특히 일반 여론조사, 마케팅 조사가 가장 활발하게 이루어지고 있는 미국의 경우에는 전통적 확률표집들에 대한 집착 때문에 할당 표집에 의존할 수밖에 없는 모바일조사에 대한 논의가 상대적으로 적은 편이다. 이것은 모바일 전화의 보급이 충분히 이루어지지 않았기 때문으로 보이며 모바일 전화 보급이 증가하게 되면 이 분야에 대한 많은 연구가 이루어질 것으로 기대된다.<sup>4)</sup> 반면 인터넷 조사방법에서는 많은 연구가 이루어지고 있으며,<sup>5)</sup> Harris Interactive의 경우는 성향지수(propensity score)를 이용한 가중치를 2000년 미국선거 예측조사에서 활용하여 성공을 거두기도 했다. 이들의 조사는 미리 모집된 인터넷 패널을 대상으로 실시되었는데, 미리 개발해 놓은 성향지수를 이용하여 가중한 결과 실제 선거결과와 가장 근접하게 예측하였다(Taylor et al. 2001, Taylor 2000).

3) 예컨대 강남준 (2002)은 모바일을 이용한 시청률과 기존 시청률 조사기관 자료를 비교하였다.

4) 2002년 AAPOR 57차 연례학술대회에서 “Mobile phones in Telephone Surveys”라는 집담회(roundtables)가 처음으로 열려 모바일 조사에 대한 학술적 관심이 커지기 시작했다.

5) 인터넷 조사방법에 관한 연구들은 Couper(2000) 참조.

〈표 1〉 가중 전/후 후보 간 지지율 변화 : Harris Interactive

	Bush	Gore	차이
가중 전 원자료	51%	42%	9%
인구통계학 가중	49%	45%	4%
성향변인 가중	47%	47%	0%
실제 결과	47.9%	48.4%	0.5%

그러나 인터넷 조사방법도 아직 명확하게 정립되어 있지는 않다. Harris Interactive의 경우 비록 선거예측에서는 성공했지만 이것이 과연 어느 정도 일반성을 가지는지는 아직 검증되지 않았다. 그리고 이론적인 측면에서도 미토프스키(Mitofsky 1999)와 테일러와 테라니언(Taylor and Terhanian 1999) 간의 논쟁에서 볼 수 있듯이 아직도 논란의 여지가 많다.

한국의 경우도 인터넷 조사의 표집방법에 대한 연구가 이루어지고는 있지만(이계오 2000, 김영원과 변종석 2000, 윤은성과 김영원 2002), 아직 Harris Interactive와 같이 성향가중 방법에 의한 비표집 오차의 해결방안을 제시하지는 못하고 있다.

따라서 방법론적 관점에서 현재 우리가 처해 있는 여론조사 환경은 기존 유선전화조사 방법에 대한 대안적 방법의 도입을 요구하고 있다. 3,200만 대에 달하는 휴대전화 보급률, 세계 유례가 없는 초고속 인터넷 통신망의 확산 등으로 기존 유선전화는 이제 낡은 통신수단이 되었다. 특히 20대 젊은 세대에게 유선전화에 의한 여론조사는 낮은 접촉률로 인해 의미 있는 결과를 얻어내기 힘들다. 그러나 사회적 커뮤니케이션 수단에서 유선전화의 차지하는 비중이 감소하고 있지만, 아직 이를 대처할 만한 조사방법이 정립되지 않고 있다. 이러한 추세가 지속된다면 사회조사가 방법론적 위기에 직면할 수도 있게 된다. 현재 전화조사가 채택률 차이 및 높은 거절률로<sup>6)</sup> 인해 조사방법의 타당성이 감소하고 있다면, 모바일과 인터넷 조사는 앞서 지적하였듯이 표본 대표성의 문제가 아직 해결되

6) 선거조사의 경우 응답성공률이 조사방법 및 회사에 따라 다르지만 15~25% 정도일 것으로 추정된다.

지 않은 상태다. 이에 따라 방송사 선거예측에서도 전화 조사, 인터넷 조사, 모바일 조사 등 다양한 조사방법 중에서 결국 출구 조사(exit poll) 방법을 택하고 있는 실정이다.<sup>7)</sup> 출구 조사란 선거예측에서만 적용할 수 있는 방법이라고 볼 때 보편적인 사회조사 방법의 정립이 시급히 요청된다.

이러한 맥락에서 본 연구는 우선 모바일 조사에서 현재 사용하고 있는 표집절차의 편향정도를 검토해 보았다. 선거조사 및 과학적 이슈에 대한 조사결과를 분석한 결과 모바일 표본조사는 센서스 자료와 비교해 볼 때 표본 편향이 나타났고, 이러한 표본 편향으로 인해 조사결과에 오차가 발생하고 있다. 그러나 여러 인구사회학적 변인을 이용한 성향지수를 산출한 다음 이를 가중치로 사용했을 때, 이러한 오차는 감소하였다. 따라서 앞으로 패널모집 절차의 개선 및 성향지수의 정교화를 통해서 보다 대표성 있는 사회조사가 가능할 것으로 전망된다.

#### 나. 모바일 조사와 다른 조사방법의 차이

모바일을 이용한 조사의 좋은 점으로는 두 가지를 들 수 있다. 첫째로 응답자 단위의 표집이 가능하다는 점이다. 방문면접 조사나 유선전화의 경우 최종 표집단위인 개인을 직접 추출하지 못하고 중간 표집단위인 가구를 먼저 추출한 다음, 가구에서 다시 개인을 추출하게 된다. 가구까지 확률적으로 표집이 되었다고 하더라도 최종 응답자 추출과정은 확률적으로 이루어지기 힘들다(양승목 등 1991 ; 조성겸 1990). 또 가구를 표집단위로 하기 때문에 조사대상자 추출이 결국 재택자 중심으로 조사가 이루어지게 된다. 이러한 문제를 극복하기 위한 방법들이 제시되고는 있지만(조성겸 1997), 조사실행과정에서 적용하기가 쉽지는 않다. 특히 한국의 경우 부재자에 대한 재접촉 시도가 그다지 효과적이지 않았다.<sup>8)</sup> 모바일

7) 본격적인 출구 조사는 2000년 총선에서 KBS와 SBS가 시작한 이래 예측력이 전화 조사보다 높은 것으로 판단되어 그 이후 2002년 지방선거, 보궐선거 및 대통령 선거의 예측조사 방법으로 계속 이용되고 있다.

8) 재접촉의 효과에 실험은 코리아 리서치센터에서 1997년 11월 세 차례에 걸쳐 실시하였는데, 평균 3회 이상의 재접촉을 시도하여 조사성공률을 23% 정도 얻

조사는 응답자 개인을 직접 표집하기 때문에 이러한 유선전화 조사에서 발생하는 문제가 나타나지 않는다.

둘째로 모바일 조사는 인터넷조사와는 달리 도달범위가 넓다. 인터넷 사용률이 연령이나 직업 등에 따라 편향이 많다면 모바일 사용 정도는 그 편향이 훨씬 덜하기 때문이다. 이외에도 모바일 조사는 응답자에 대한 정보획득이 용이하다는 점, 모바일 플래쉬(mobile flash) 등의 기술개발로 인해 그림이나 음성, 동영상을 이용한 조사가 가능하다는 점 등의 장점이 있다.

이러한 모바일 조사는 표본 프레임 구성방식에 따라 두 가지로 구분될 수 있다. 하나는 모바일 전화번호를 무작위로 추출하여 표집하는 방법이고 또 하나는 미리 모집된 패널에서 표본을 추출하는 방법이다. 전화번호를 무작위로 추출하거나 발생시켜서 조사하는 경우 확률표집이 이루어질 수 있기 때문에 이상적이기는 하지만 모바일 전화 특성상 거절 또는 비수신 등으로 조사성공률이 낮을 수밖에 없고 이로 인해 확률표집으로서의 특성이 사라진다.<sup>9)</sup> 미디어 리서치에서 2002년 10월 24~25일에 실시한 실험조사에 의하면 RDD방식으로 추출한 모바일 전화번호 중에서 조사에 성공한 것은 9.9%에 지나지 않았다. 이러한 응답성공률로는 확률표집의 특성이 없다고 보아야 한다. 예컨대 응답자의 성별분포의 경우를 보면 남자가 70.1%, 여자가 29.9%였다. 이번 실험에서 조사내용은 12월 대통령 선거에서의 지지후보에 관한 것이었다. 만약에 일반인들의 관심도가 높지 않은 문제의 경우에는 응답률이 훨씬 낮을 것으로 예상된다. 따라서 이러한 RDD방식에 의한 모바일 조사는 비록 이론적으로는 우수하지만 실제 사회조사 방법으로 사용하기에는 한계가 있다.

두 번째로는 조사 대상자가 될 의사가 있는 사람들을 모집하여 미리 패널을 구성한 다음 이러한 패널에서 일정 수의 표본을 추출해서 조사하는 패널 조사방법이다. 이러한 패널 조사방법은 희망자를 대상으로 했기 때문에 거절을 문제가 발생하지는 않는다. 그렇지만 패널에서 추출된 표본

---

었지만 조사결과에서 할당표집과 유의미한 차이가 나타나지 않았다.

- 9) RDD방식을 사용하는 모바일 조사의 또 다른 문제점은 현재 정보통신법상 휴대전화에 대한 무작위 호출은 스팸메일로 규정되어 제한을 받게 되는 점에 있다.



이 일반성을 얼마나 가질 수 있는지가 문제가 된다. 패널 참여자는 전국민을 대상으로 확률적으로 추출된 것이 아니라 인터넷이나 전화 등을 이용해 희망자를 모집해서 구성하기 때문에 패널 참여의 기회가 동일하게 제공되지 않는다. 전화로 모집할 경우에는 채택률에 따른 편향이 그리고 인터넷 모집의 경우에는 인터넷 사용 정도에 따른 편향이 패널에 나타날 수밖에 없다. 이러한 이유에서 모바일 전문조사회사인 엠비존이 구축한 100만 명 규모의 패널도 성, 연령, 학력 등의 분포에서 편향이 나타난다.

이 두 표집방법에 대한 비교는 모바일 조사방법을 사용한 경우는 아니지만 미국 2000년 대선에서 인터넷 조사방법을 사용한 Knowledge Network와 Harris Interactive의 결과를 살펴보면 예측할 수 있다. 확률적 표본 추출방법에 충실하게 인터넷 패널을 구축한 예가 Knowledge Network의 조사패널이고 비확률적 표집기법으로 패널을 구축한 예가 Harris-Interactive의 예이다. Knowledge Network(이하 KN)는 확률적 표집틀 방식을 채택해 패널을 구축하였다. RDD방법으로 미국 전역에 대한 대표성 있는 전화번호 표본을 선정한 후 역으로 이 번호의 거주자 주소를 알아내어 편지를 보내거나 직접 전화를 걸어 인터넷 조사패널에 참가해 달라고 권유해 패널을 구축하였다(Krotki & Dennis, 2001). 따라서 KN는 전집(total residential telephone household)이 확실하게 규정되어 있어 확률적 표본이라고 볼 수 있고, 이에 따라 계층별 비협조자에 대한 비율을 전집과 비교해 추정할 수 있어 확률적으로 타당한 가중방법을 만들 수 있다. 반면, Harris는 자체 웹이나 Exite.com 웹을 통해 자발적인 패널 참여자를 모집한다. 따라서 비확률적 표본이라고 볼 수 있다. 현재 전세계적으로 600만이 넘는 패널이 등록했다고 자랑한다. 자발적인 패널 등록을 할 때 기본적인 인구통계학적 변인을 수집해 이것을 추후 실제 조사에서 사용하는 표본을 구성하는데 할당자료(quota reference)로 사용한다.

Knowledge Network와 Harris Interactive의 2000년 미국 대선 예측결과는 <표 2>와 같이 일단 Harris의 판정승으로 나타났다. Harris가 더 정확한 예측을 할 수 있었던 것은 패널 구축 방법의 차이이기보다는 앞에서 설명한 새로운 가중방법(성향추정)의 개발에 따른 결과라고 본다. 그러나 Harris가 사용한 성향추정방법은 전형적인 데이터 의존적 추정방식을 취

하고 있으므로 다음 선거 예측에서는 어떠한 결과가 나타날지 모른다. 적어도 일정기간에 걸쳐 충분한 비교결과(tracking record)가 쌓이지 않는 한 어느 패널구축 방법이 더 우수하다고 볼 수 없다. 테일러(Taylor et. al. 2001)도 Harris의 2000년 선거예측 정확도는 놀랄 만하지만 이번에 사용한 성향추정 기법이 2년 후 중간 선거나 4년 후 대선 때에도 그대로 맞아떨어지리라는 보장은 없다고 하였다.

〈표 2〉 인터넷 조사패널에 따른 후보간 지지율 비교

	Gore	Bush	차이
실제 결과	48.4%	47.9%	0.5%
Knowledge Net.	44	46	2
Harris Inter.	47	47	0

현재 우리나라에서 사용하고 있는 모바일조사 표집방법은 거의 대부분 패널 조사방법으로 이루어지고 있다. 확률표집 방식에 의한 조사자료는 현재 비교가 가능할 만큼 축적되어 있지 않다. 따라서 이 연구에서도 패널조사방법으로 한정하여 모바일 조사방법을 살펴보겠다.

#### 다. 할당표집과 표집오차

패널을 이용할 경우 발생하는 편향을 제거하기 위해 현재 모바일 조사에서는 성, 연령, 지역을 기준으로 할당표집을 실시한다. 이러한 할당표집 방법은 표본 편향을 제거하는 방법으로 모바일 조사뿐만 아니라 전화, 인터넷, 면대면 조사 등에서도 흔히 이용된다. 할당표집을 하게 되면 최소한 할당기준이 된 변인에서는 전집과 동일한 분포를 갖도록 표본을 구성할 수 있는데 그 기준으로는 전집분포를 센서스자료 등에서 알 수 있는 성별, 연령, 거주 지역이 사용 된다.<sup>10)</sup>

10) 조사회사에 따라서는 연령 할당을 하지 않고 사후에 가중치를 이용하여 보정하는 방법을 사용한다. 한국의 경우 교육수준이나 소득 등은 할당변인으로 거의 사용되지 않는다.

그러면 이러한 할당표집이 표본 편향을 제거해 줄 수 있는가? 할당표집 방법으로 추출된 표본은 대표성 정도를 알 수 없다. 그럼에도 불구하고 할당표집으로서 충분히 대표성을 가지고 있다고 믿는 경우도 있다. 이러한 할당표집에 대한 믿음에서 또는 때로는 할당표집의 문제점을 인식하면서도 현실적인 이유에서 할당표집은 한국에서 널리 이용된다. 확률표집 방법을 사용하게 되면 할당표집 방법에 비해 경비가 많이 소요될 뿐만 아니라<sup>11)</sup> 당일조사가 불가능하게 된다.

할당표집 방법을 사용할 경우에 어떤 오차가 발생할 수 있는가? 면접원에 의한 면대면 조사에서 할당표집방법을 사용하게 되면 면접원이 접촉하기 쉬운 응답자를 선택하기 때문에 발생하는 선정오차가 발생한다(Marsh and Scarbrough 1990). 예컨대 서베이 조사에서 고학력층이 상대적으로 높게 나타나는 것은 이러한 선정오차에서 비롯되었을 가능성이 높다. 면접원들이 고학력자로부터 응답을 받기가 더 쉽기 때문이다. 이러한 면접원에 의한 선정오차가 영국 총선예측의 오차요인으로 작용했다는 분석결과도 제시되었다(Broughton 1995).

전화조사의 경우는 면접원의 자의적 판단에 의한 선정오차는 개입될 소지가 적다. 전화번호 자체를 컴퓨터로 추출해서 제공해 주기 때문이다. 그렇지만 유선전화를 통해 접촉할 수 있는 대상자는 전화조사가 이루어지는 시간대에 집에 있는 재택자로 한정된다. 통상 전화조사가 밤 9시 30분까지 이루어진다고 볼 때 이 시간 이후에 집에 들어오는 사람들은 전화조사에서 제외된다. 이 시간 이후 귀가자는 직업이나 학력 등에서 차이가 있기 때문에<sup>12)</sup> 조사결과는 편향될 수밖에 없다.

모바일의 경우는 패널에 포함된 사람 중에서 추출하기 때문에 재택물이나 면접원과는 무관하다. 다만 패널 자체가 전집과 다를 경우에 편향이 나타날 수 있다. 즉 패널과 패널에 포함되지 않은 사람들 간에 성이나 연령 또는 거주지역 외에 다른 부분 예컨대 교육수준이나 인터넷 사용정도

11) 2002년 5월에 실시한 미디어 리서치 실험조사에서 3회 재접촉 확률표집방법을 사용할 경우 할당표집을 사용할 경우보다 세 배의 경비가 소요되었다.

12) KBS와 서울대 언론정보연구소에서 1981년부터 실시하고 있는 국민생활시간 조사 보고서 참조.

등에서 체계적인 차이가 발생하게 되면 표본의 대표성이 떨어지는 것이다. 현재 패널 모집이 인터넷에 의존하는 경우가 많기 때문에 이러한 오차요인을 무시하기는 어렵다. 김준한(2000)의 연구에서 동일 주제를 가지고 인터넷과 전화조사를 하여 성, 연령, 지역이 동일한 집단끼리 비교해도 인터넷상의 응답자와 전화 응답자 간에 차이가 나타났다. 이 점은 인터넷으로 모집된 패널이 할당표집을 한다 하더라도 전집의 대표성을 띠기 어렵다는 것을 뜻한다.

지금까지 살펴본 바와 같이 할당표집 방법은 면대면 조사나 전화조사 상황은 물론 모바일 상황에서도 이론적으로나 실제적으로 표집방법이 갖는 문제점을 해결해 주지는 못한다. 따라서 이러한 편향요인을 정확히 분석하고 그것을 보정하거나 감소시킬 수 있도록 표집이나 분석절차를 개선할 필요성이 있다.

한국의 경우 1996년 총선, 2000년 총선과 같이 선거예측이 크게 빗나가는 등 표집방법에 대한 검토가 요구되는 상황이 많았음에도 불구하고, 아직까지 할당표집에 대한 대안이 도입되지 않았다. 이것은 우선 조사결과가 빗나갔을 때 과연 이 오차 중 얼마가 조사표집의 문제에서 비롯되었는지 구별하기 쉽지 않기 때문일 수도 있다. 표본의 편향 외에도 면접원의 면접태도, 질문내용, 사회적 분위기, 조사시기와 투표시기의 차이, 투표율 등이 모두 예측오차와 관계가 있는 것으로 보이기 때문이다(박용치 2002 ; 홍내리와 허명희 2001 ; 조성겸 2001). 그렇기 때문에 전화조사의 오차가 발생했을 때 그것이 표집절차상의 문제인지 아니면 다른 요인 때문인지 분간이 어렵다. 한국의 경우는 조사오차의 원인으로 사회적 분위기(이준웅 2001)나 응답자의 거짓 응답(류제복 2000) 등이 지적되는 경향이 많지만 실제 분석결과 이러한 요인보다는 표집방안일 가능성이 높다는 연구결과도 있다(조성겸 2001). 조사실무자들의 경우 이러한 경향이 강하다. 할당표집이 가장 경제적인 조사방법이라는 요인 외에도 이러한 학계 및 조사 실무자들의 인식이 표집절차에 대한 반성을 어렵게 만드는 요인이 되고 있다. 두 차례에 걸친 예측실패에도 불구하고 표집의 절차나 방법에 변화가 없다. 이 점은 영국이 1992년 예측 실패 이후 조사방법에 대한 다양한 변화를 시도하는 것과는 대비된다.

할당표집에서의 문제점을 극복하는 방법은 두 가지 방향이 있다(강남준 2002). 하나는 확률표집 원칙을 가급적 준수하도록 하는 것이다. 예컨대 패널에 의존하기보다는 직접 전화번호를 추출하는 조사방법을 사용하면 거절률을 줄여나가는 방법을 들 수 있다. 또 하나는 성향지수 등 가중치를 이용해 감소시키는 방안이다. Harris Interactive의 방법이 여기에 속한다. 본 연구는 우선 기존의 인구학적 속성을 이용한 성향지수를 적용해 봄으로써 과연 이러한 노력이 가능한지를 살펴보았다.

가중치를 이용해서 표본의 특성을 보정할 경우 그 효과가 언제나 긍정적이지는 않다. 가중치에 사용된 변인이 조사하고자 하는 내용과 관련이 없는 경우에는 조사결과에 영향을 주지 않거나 때로는 더 나쁘게 만들 수도 있기 때문이다. 베호바(Vehovar 1999)가 슬로베니아에서 인터넷 조사와 전화 조사를 실시하고, 성, 연령, 인터넷 접촉 관련 행동들을 이용해서 가중치를 부여했지만 가중치가 조사결과에 큰 변화를 주지 않았다. 베스트와 크로거(Best & Krueger 2000)의 연구에서는 가중치 부여가 오히려 결과를 더 나쁘게 만들었다. 즉 표본에 특정 집단이 체계적으로 배제되었을 경우에는 가중치를 부여한다고 해서 그것이 해결되지 않는다는 것을 의미한다. 예컨대 선거예측조사에서 20대의 민주당 지지자들이 조사에서 누락되었고 한나라당 지지자들이 조사에 포함되었을 경우 연령을 기준으로 가중치를 부여하면 한나라당 지지도가 더 높게 나타날 것이다.

가중치로 표집의 문제를 해결하기 위해서는 누락된 집단과 그렇지 않은 집단을 구분해 줄 수 있는 속성이 가중치에 포함되어야 할 것이다. 따라서 만약에 성별이나 연령, 학력 이외에 일반적인 사회적 태도나 인식에 관련되었을 다수의 요인들을 가중치에 반영할 수 있다면, 어느 정도 이러한 편향을 감소시킬 수 있을 것으로 보인다. 통계청에서 이루어지는 다양한 사회지표들을 가중치로 활용할 경우 표집의 편향을 감소시킬 가능성도 있다.

#### 라. 연구문제

모바일 조사가 새롭게 사용되고 있지만 그 표본의 대표성이 어느 정도

일지는 앞서 살펴본 바와 같이 명확치 않다. 모바일이 의존하고 있는 표집절차 그 자체로서는 대표성을 충분히 보장하지 않기 때문이다. 본 연구는 우선 모바일 조사의 현행 표집체제에서 조사의 표본이 어떤 측면에서 어느 정도의 편향을 보이며 이것이 조사결과에 어느 정도 영향을 주는지 살펴보고자 한다.

둘째는 이러한 편향을 가중치를 이용한 보정 방법을 통해 어느 정도 감소시킬 수 있는지를 검토해 보았다. 흔히 표본상의 대표성 부족을 가중치를 이용해서 보정하는 방법이 많이 사용되고 있다. 따라서 연령, 성, 지역적 할당표집을 실시하게 되더라도 여전히 나타나게 되는, 예컨대 교육 정도의 편파 같은 부분을 가중치를 이용해서 보정하게 될 경우 과연 오차가 감소될 것인지를 살펴보고자 한다.

## II. 연구자료와 분석방법

### 가. 분석자료

모바일로 실시된 조사자료는 많다. 그렇지만 표본의 편향정도 및 그것의 영향을 살펴보기 위해서는 모치수를 알 수 있거나 또는 다른 조사방법과 동일하게 시행된 것이어야 한다. 이러한 면에서 본 연구는 엠비존에서 실시한 6·13 지방선거 예측 조사자료와 한국의 엘시(ELSI) 연구팀에서 2002년 4월과 10월에 실시한 전화와 모바일 조사자료를 분석자료로 선택하였다.<sup>13)</sup> 엠비존의 선거자료의 경우는 표본의 편파는 물론 조사결과와 정확도도 같이 알아볼 수 있다. 그러나 배경변인이 성, 연령, 거주지역, 교육수준, 직업 등으로 제한되어 있다. 배경변인을 이용한 가중치 부여의 효과를 알아보기 위해서 엘시연구팀에서 실시하는 사회적 인식조사

13) 엘시(ELSI)연구팀은 사회학, 법학, 철학, 커뮤니케이션학, 교육학 등 다양한 분야의 연구자들이 참여하여 생명공학에 대한 윤리적·법률적·사회적 함의를 연구하고 있으며, 그 일환으로 생명공학에 관한 다양한 사회 인식조사를 실시해 오고 있다.

와 동일한 조사를 모바일로 실시하면서 여러 가지 배경변인들을 포함시켰다.

모바일 표본이 어떤 편향을 갖는지는 2000년에 실시된 센서스 데이터와 비교해서 알아보았다. 그리고 이러한 편향이 조사결과에 미치는 영향은 투표결과 예측의 오차를 통해 살펴보았다. 엠비존에서 실시한 지방선거 자료는 6·13 선거 당일에 유권자를 대상으로 서울과 경기 두 지역에서 실시되었는데, 97만 명으로 구성된 패널 중에서 성, 연령, 지역별 할당표집 방법을 이용하여 각각 1,000명 규모로 조사되었다. 이 외에도 선거예측을 위한 모바일 조사회사인 NI Korea에서 실시한 것도 있지만 이 조사의 경우 미리 구성된 패널에서 표집한 것이 아니라 선거조사를 위해 패널을 구성한 것이기 때문에 제외했다.

본 연구에서는 또한 엘시 연구팀이 전화로 실시한 사회여론조사 자료와 이것과 동시에 엠비존에서 실시한 모바일 조사자료도 비교분석하였다. 엠비존 모바일 조사는 전화조사와의 비교를 위해 동일한 설문지를 이용하여 실시한 것이다. 다만 배경변인은 10월 조사에서만 추가되었는데, 4월 조사를 실시할 당시에는 배경변인을 이용한 가중치 부여방법이 고려되지 않았기 때문이다.

#### 나. 성향지수(propensity score)를 이용한 가중치 산출방법

본 연구는 표본의 편향을 성, 연령, 거주지역, 교육수준, 직업, 주택소유 여부, 자동차 소유 여부, 인터넷 사용 정도, 신문구독 정도, 인터넷 회선 보유 여부, 종교, 도시거주 여부 등 12개의 변인을 이용하여 보정하였다. 먼저 이들 변인을 가지고 성향지수를 구한 다음 전화 조사표본과 모바일 조사표본의 성향지수 분포가 동일하게 되도록 모바일 조사표본에 가중치를 부여하였다. 성향지수는 개인이 특정 표본에 포함될 확률을 의미하며 (Rosenbaum and Rubin 1984 ; Rosenbaum and Rubin 1983), 보통 인구학적 속성 및 상황적 요인 등을 이용하여 산출한다. 이 지수는 두 개의 세베이 데이터 또는 서베이 데이터와 데이터 베이스를 연결하여 분석할(cross-design synthesis) 경우에도 활용되고 있다(Boruch and Terhanian 1996).

테일러(Taylor, 2000 ; Taylor et. al. 2001)가 주장한 바와 같이 자발적 조사패널에 따른 표본 편파는 단순히 인구 통계학적 변인 만을 사용한 가중 방법만으로는 수정되지 않는다. 예측변인과 연관된 성향추정 변인을 사용한 가중이 필요하다. 테일러가 밝힌 대로 이러한 변인은 태도나 행동과 관련된 변인들이 주종을 이룬다. 예를 들어, 선거에서 지지하는 후보에 대한 가중 성향변인으로 어떤 것이 더 적합할 것인가에 대해 논의해 보자.

다음 <표 3>은 엘시연구에서 두 가지 서로 다른 생명관련 주제에 대한 휴대전화 표본과 일반전화 표본 사이의 차이에 관한 것이다. 낙태 찬성론자나 안락사 찬성론자 모두 생명에 대한 진보적 생각을 가진 사람들이다. 그러나 두 이슈는 그것이 가지고 있는 사회문화적 의미가 서로 다르다. 예를 들어, 낙태에 찬성하는 사람은 안락사에 찬성하는 사람보다 더 진보 성향이라고 볼 수 있다. 낙태는 태어나지 않은 생명에 대한 조치이고 안락사는 죽을 생명에 대한 판단이라고 볼 수 있다. 따라서 미국에서 일반적으로 아주 진보성향의 정치가나 사회운동가들이 낙태 찬성론자(pro-choice)인 경우가 많다. <표 3>에 의하면 낙태의 경우 ‘경우에 따라’ 비율이 두 표본 모두 55% 정도이고 안락사의 경우도 24%~29%로 서로 비슷하다. 그러나 찬성의 경우에는 휴대전화 표본의 응답비율이 낙태는 10%, 안락사는 17% 정도 더 크다. 따라서 두 표본의 차이를 최소화하기 위한 가중변인으로는 숫자상으로 안락사 변인이 더 적합한 성향변인이 될지도 모른다. 그러나 예측변인이 대선후보에 대한 여론조사일 경우 문제는 그렇게 간단하지 않다. 성향가중변인으로 선정되기 위해서는 예측변인과의 공변량(co-variance)이 중요하지 대상전집과 비교한 표본 간 편차의 크기는 큰 문제가 되지 않는다(D’Agostino & Rubin, 2000 ; Rubin, 1997). 따라서 지난 대선과 같이 후보들이 진보와 보수로 극명하게 갈라져 있는 상태에서 여론조사를 할 경우 적절한 성향가중 변인은 안락사보다는 낙태가 더 적절할 것이다. 그 이유는 앞에서 논의한 것처럼 정치적인 진보성향은 안락사변인보다는 낙태변인과의 공변량이 더 크다고 볼 수 있기 때문이다.

이 연구에서는 12개 성향변인의 교차표를 구한 다음 각 셀 단위로 가중치를 부여하는 방법이 불가능하기 때문에 로지스틱 회귀분석(logistic regression) 방법을 사용하여 성향지수를 계산했다. 이렇게 계산한 성향지수를 5개 급



간으로 구분한 다음 그 빈도분포가 모바일과 전화조사 표본이 일치하도록 모바일 표본에 가중치를 부여하였다.

<표 3> 주요 이슈에 대한 의견 : 엘시 조사

이슈	태도	전화	모바일
낙태	찬성	7.3%	17.2%
	반대	33.0	27.8
	경우에 따라	55.1	54.1
	모르겠다	4.6	1.0
안락사	찬성	49.3	66.0
	반대	14.6	9.0
	경우에 따라	29.1	24.2
	모르겠다	7.0	0.9

### III. 모바일 조사의 편향 정도와 영향

그러면 모바일 조사의 표본은 어느 정도의 편향을 보이는가? 연령별, 성별, 지역별 변인은 할당표집을 하기 때문에 이 부분에서는 전집분포와 일치한다. 따라서 교육, 직업 등 할당에 사용되지 않은 특성이 센서스 자료와 어느 정도 차이가 있는지 살펴보았다.

먼저 6·13 선거의 경우를 살펴보았다. 다음 표에서 보는 바와 같이 고학력, 그리고 직업의 측면에서는 화이트칼라층이 과대표집되고 가정주부 등이 적게 뽑히는 경향을 보였다. 이것은 전화조사의 표본 편향과는 상반된 모습을 보인다. 전화조사는 동일지역에서 투표 전날 KRC와 MRI가 성, 연령, 거주지역에 입각한 할당표집 방법으로 실시한 것이다. 이 전화조사의 경우는 가정주부나 무직자 등이 과대표집된 반면, 화이트칼라와 블루칼라 층은 과소추출되었다. 학력에서는 중졸 이하의 경우가 전화와 모바일 모두에서 과소표집되고 있지만 특히 모바일 조사에서 더 심했다.

〈표 4〉 선거예측조사에서 모바일 조사자료의 편향 정도  
(성, 연령, 거주지역 할당표집)

지역	변인	센서스 (통계청)	전화 (MRI/KRC)	모바일 (엠비존)	인터넷** 패널조사	
서울	표본크기	137,874명	1,503명	923명	8,061명	
	직업	화이트칼라	26.8%	15.0%	33.0%	44.7%
		블루칼라	13.8	9.9	15.2	18.9
		자영업	13.7	22.7	13.3	8.3
		주부	28.9	35.0	19.4	18.3
		학생	7.6	7.7	8.7	9.7
		기타/무직	8.6	9.8	10.4	0.1
		X <sup>2</sup> 검증		136.7*	52.5*	288.4*
	학력	중졸 이하	22.8	15.5	7.3	1.8
		고졸	37.1	40.4	41.0	27.6
대재 이상		40.1	44.2	51.7	70.7	
X <sup>2</sup> 검증			30.8*	143.0*	451.3*	
경기	표본크기	116,640명	1,603명	1,026명	5,468명	
	직업	화이트칼라	24.0%	18.2%	33.5%	39.2%
		블루칼라	16.3	7.2	18.9	20.5
		자영업	14.5	19.2	14.6	9.2
		주부	29.4	37.4	17.8	22.3
		학생	4.9	6.0	5.6	7.5
		기타/무직	10.9	12.0	9.6	1.3
		X <sup>2</sup> 검증		105.4*	91.2*	241.9*
	학력	중졸 이하	26.7	19.6	6.5	3.0
		고졸	42.5	39.4	53.8	35.0
대재 이상		30.8	41.1	39.7	62.0	
X <sup>2</sup> 검증			55.6*	208.6*	539.7*	

\* p<0.001

\* 일부 지역이나 연령층에서 할당된 수만큼 표집되는 않는 경우가 있었다. 이 경우는 가중치를 이용해서 보정하였다.

\*\* 인터넷 조사회사에서 같은 시기에 실시한 패널 조사결과.

<표 5> 엘시 사회조사에서 모바일 조사자료의 편향 정도

		센서스	10월조사(목, 금, 토, 일)		4월조사(금, 토, 일)	
			전화	모바일	전화	모바일
표본크기		606,851명	1,100명	930명	1,000명	1,007명
직업	화이트칼라	24.2%	19.5%	18.7%	22.0%	33.0%
	블루칼라	16.2	12.4	23.4	11.4	18.5
	자영업	12.0	11.4	12.6	10.3	11.7
	농,임,어업	7.8	7.7	0.2	6.8	0.7
	가정주부	26.8	24.9	18.0	22.8	15.3
	학생	5.5	12.0	12.8	10.6	8.4
	무직/기타	7.5	12.1	14.5	16.1	12.5
	$X^2$ 검증		124.7*	310.0*	171.8*	197.9*
학력	중졸 이하	35.1	26.4	7.6	20.4	12.8
	고교 졸업	33.1	31.8	41.8	34.0	42.7
	대학 재학	7.6	12.0	8.7	13.6	14.0
	대졸 이상	24.2	29.8	41.9	32.0	30.5
	$X^2$ 검증		60.5*	369.4*	139.4*	239.8*

\* p<0.001

또 <표 5>에서 보듯이 엘시 조사에서도 여전히 편향이 나타났다. 다만 전화 조사에서 4월 조사가 10월 조사보다 편향이 많았던 것은 4월은 목요일부터 일요일까지, 10월은 금, 토, 일에 이루어졌기 때문으로 보인다. 즉 휴일에 이루어진 조사가 덜 왜곡되는 경향을 보이는 것으로 해석될 수 있다.

이러한 표본의 편향은 직업과 교육뿐만 아니라 다른 인구통계학적 속성에서도 나타나고 있다. 인터넷 회선 사용정도, 혼인상태, 종교, 주택소유 여부, 차량 소유 정도 등에서 많은 차이가 나타나고 있다. 이러한 점으로 볼 때, 전화나 모바일 모두 대표적인 표본을 추출한다고 보기는 어렵다.

〈표 6〉 기타 속성에서의 편향(엘시 10월 조사)

속 성		센서스	전화	모바일	속 성		센서스	전화	모바일
인터넷	매일 사용	20.6%	42.0%	82.5%	신문구독	매일 읽는다	53.0%	54.8%	57.5%
	가끔 사용	14.7	19.2	15.8		가끔 읽는다	11.8	19.4	37.4
	사용 안 함	64.7	38.8	1.7		읽지 않는다	35.2	25.8	5.2
	$X^2$ 검증	117.4* 614.3*				$X^2$ 검증	74.1* 811.0*		
회선	있다	78.6	72.6	89.3	종교	불교	27.1	25.5	19.0
	없다	21.4	27.4	10.7		기독교	19.2	25.0	30.3
	$X^2$ 검증	21.4* 68.1*				천주교/기타	7.3	9.8	16.3
혼인	미혼	19.5	27.7	34.0		종교 없음	46.4	39.7	34.5
	기혼	80.5	72.3	66.0		$X^2$ 검증	17.2* 140.5*		
	$X^2$ 검증	42.8* 133.9*				차량	없다	37.3	26.5
주택	소유	54.3	75.7	69.2	1대		47.4	55.5	57.2
	비소유	45.7	24.3	30.8	2대 이상		15.3	18.1	23.1
	$X^2$ 검증	184.5* 89.5*			$X^2$ 검증		19.0* 60.0*		

\* p&lt;0.001

이러한 분석결과에서 다음과 같은 점들을 알 수 있다. 하나는 같은 표집절차를 사용하고 있지만 표본의 특성은 다르다는 점이다. 편향의 내용이나 정도에서 모바일과 전화가 다르다. 따라서 오차의 크기나 방향도 다르게 된다. 둘째는 이러한 편향이 일반적으로 측정되는 직업과 학력에서만 나타나는 것이 아니라 다양한 부분에서 나타나고 있다는 점이다. 따라서 편향의 영향은 사회조사결과 전반에 영향을 미칠 수 있다는 점이다. 셋째로 편향의 크기나 정도를 명확하게 예측하기 어렵다는 점이다. 따라서 편향이 나타나는 정도를 감안해 자료를 해석하는 것도 쉽지 않을 것으로 보인다.

그 다음으로는 표본 편향이 조사결과에 어떻게 영향을 미쳤는지를 살펴 보았다. 우선 엘시 조사의 경우를 보면 다음 표에서 보는 바와 같이 중요한 조사문항에서 각기 다른 응답분포가 나타나고 있다.

<표 7> 전화와 모바일 조사 간의 결과차이(10월 엘시 사회인식조사)

질 문	응 답	전화	모바일	X <sup>2</sup> 검증
친자 확인, 미아찾기, 암이나 심장질환 진단 등과 같은 유전자 정보 관련 문제에 관심이 있으십니까?	관심이 있다	58.1%	70.5%	33.8 p<.001
	관심이 없다	41.9	29.5	
만약 귀하가 원하는 정보를 알 수 있다면 귀하는 유전자 검사를 받을 의향이 있습니까?	받겠다	65.2	82.8	85.2 p<.001
	받을 생각이 없다	25.6	14.4	
	잘 모르겠다	9.2	2.8	
전반적으로 유전자 검사가 사회에 미치는 영향은 어떤 것이라고 생각하십니까?	불이익이 크다	12.3	21.2	110.8 p<.001
	이익이 크다	59.9	68.7	
	잘 모르겠다	27.8	10.1	

<표 8> 전화와 모바일 조사 간의 결과 차이(4월 엘시 사회인식조사)

질 문	응 답	전화	모바일	X <sup>2</sup> 검증
귀하는 배아복제에 대해서 들어보 셧습니까?	들어보았고 대략 알고 있다	45.7%	58.7%	96.4 p<.001
	들어보았지만 잘 모르겠다	29.8	32.9	
	들어본 적도 없다	24.5	8.5	
배아복제를 어느 정도 허용해야 한다고 생각하십니까?	전면 금지해야 한다	9.0	12.8	81.3 p<.001
	난치병 치료 목적에 한해 허용해야 한다	59.2	68.6	
	연구 목적까지 허용해야 한다	22.4	17.2	
	전면 허용해야 한다	2.3	0.6	
	모르겠다	7.2	0.8	
인간복제를 허용해야 한다고 생각하십니까? 아니면 허용하지 말아야 한다고 생각하십니까?	허용해야 한다	9.3	12.8	38.5 p<.001
	허용하지 말아야 한다	71.2	77.2	
	잘 모르겠다	19.5	10.0	

두 조사 모두에서 생명공학에 대해 허용적인 태도가 모바일 조사에서 더 강하게 나타났다. 그렇지만 어느 것이 더 전집분포와 유사한 것인지는

두 표본 모두 편향되어 있기 때문에 판단하기 어렵다. 단순히 할당표집으로 여론을 조사하게 되면, 이처럼 조사방법이 무엇이냐에 따라 사회적 여론을 다르게 파악하게 된다.

한편, 선거조사의 경우는 조사방법간의 차이가 엘시 사회조사의 경우보다는 적게 나타났다. <표 9>에서 보듯이 경기와 서울에서 모바일의 예측 오차가 조금 크게 나타났다. 전화조사는 두 군데 모두 오차범위내에 있다. 그렇지만 조사방법간의 차이는 엘시 조사 경우 보다 훨씬 적었다. 이것은 아마도 우리나라 투표성향이 연령이나 지역적 특성의 영향을 많이 받기 때문인 것으로 해석된다. 즉 할당기준이 되는 연령과 거주지역이 투표성향과 밀접한 관련이 있기 때문에 이 두 부분의 분포를 전집과 동일하게 함으로써 어느 정도의 정확성은 기할 수 있는 것이다. 그렇지만 엘시조사에서 조사한 생명공학에 대한 태도는 이러한 연령이나 지역적 성향과 관련성이 적기 때문에 조사방법간 차이가 크게 나타났다는 것이다.

<표 9> 모바일 조사를 이용한 6·13 선거예측결과

지역	실 제		모바일		전 화	
	순위	득표율	지지도	오차	지지도	오차
서울	1위	52.3%	53.9%	1.6%	51.9%	0.4%
	2위	43.0	42.0	1.0	45.8	2.8
경기	1위	58.4	55.2	3.2	57.3	1.1
	2위	36.0	39.4	3.4	39.6	3.6

#### IV. 가중치의 효과

표본의 편향을 가중치를 사용해서 보정하게 되면 오차가 감소할 수 있는가? 가중치를 이용해서 표본의 편향을 보정하는 방법이 일반적으로 사용되고 있다. 그렇다면 이러한 방법을 보다 적극적으로 사용해서 모바일 조사의 표집상의 문제를 해결할 수 있는가?

가. 직업과 학력 보정

성, 연령, 거주 지역의 경우 할당비율로 전집과 분포를 맞추게 된다. 따라서 표본의 편향은 대체로 소득과 학력으로 파악한다. 소득의 경우 무응답자도 많고 측정자체에 대한 정확도가 약하다. 또 측정방법도 상이하기 때문에 가중치로 사용하기에 어려움이 있다. 본적지의 경우 선거조사에서는 유용하지만 사회조사 일반에 적용시키기는 문제가 있다. 따라서 본 연구는 보다 보편적으로 사용될 수 있고 대부분의 사회조사에서 측정하는 직업과 학력을 이용하여 성향지수를 구한 다음 가중치를 구해 보았다. 그리고 여기에 소득수준이나 사회적 위치를 나타낼 수 있다고 보는 성, 연령, 지역, 직업, 학력 등 다섯 개의 변인을 추가해서 가중치를 구해 보았다. 마지막으로 앞에서 언급한 12개의 관련 성향변인을 모두 사용해 결과를 보정하였다.

<표 10> 직업과 학력을 보정한 경우의 예측오차 변화

		전화		모바일		인터넷	
		1위	2위	1위	2위	1위	2위
서울	실제	52.3%	43.0%	52.3%	43.0%	52.3%	43.0%
	예측치	51.9	45.8	53.9	42.0	51.1	45.6
	보정*	52.1	44.9	53.0	43.4	44.2	52.6
	오차 변화	-0.2	-0.9	-0.9	-0.6	+6.9	+7.0
경기	실제	58.4	36.0	58.4	36.0		
	예측치	57.3	39.6	55.2	39.4		
	보정*	58.3	38.4	55.1	40.0		
	오차 변화	-1.0	-1.2	+0.1	+0.6		
부산	실제	63.8	19.4			63.8	19.4
	예측치	63.1	19.6			57.4	20.7
	보정*	67.1	18.9			54.9	21.9
	오차 변화	+2.6	+0.3			+2.5	+0.8

\* 성, 연령, 지역, 직업, 학력을 동시에 투입하여 센서스 데이터를 기준으로 성향지수를 구한 다음, 이것을 이용하여 가중치를 부여한 것이다.

선거조사에 직업과 학력을 이용한 가중치를 적용해 본 결과 <표 10>에서 보는 바와 같이 예측치에 큰 변화가 나타나지 않았다. 즉 보정의 효과가 없는 것이다. 그래서 추가로 부산지역 전화 조사표본을 같은 방법으로 분석해 보았다. 그 결과 여기에서는 가중치를 부여함으로써 오히려 오차가 증가하였다. 이것은 앞서 가중치를 사용한 다른 연구들의 경우와 같은 결과라고 할 수 있다. 즉 직업과 학력 이외의 부분에서 편향이 나타나고 있는 상태에서 단지 두 변인만을 가중치에 반영할 경우 오히려 더 왜곡시킬 수 있다는 점을 보여준다. 이 점은 전화나 모바일보다 편향이 조금 더 심했던 인터넷 조사표본에 적용했을 때 오차가 7% 정도나 더 증가하는 경우도 발생한다는 점에서도 알 수 있다. 이와 같은 결과는 단순히 인구학적 속성상의 편향을 제거하기 위해 가중치를 부여하는 것이 위험하다는 점을 보여준다.

#### 나. 다양한 사회경제적 변인의 보정효과

통상적인 인구통계학적 변인을 가중한 것과 다시 5개의 성향변인, 그리고 성, 연령, 지역, 직업, 학력, 종교, 혼인, 인터넷 사용 정도, 신문 구독 정도, 차량 소유 여부, 주택 소유 형태 등 12개 변인을 사용해 보정한 결과에 대한 비교는 <표 11>에 나타나 있다. 이와 같이 한 이유는 표본에서 체계적으로 누락된 집단의 특성도 이러한 12개 변인을 동시에 투입하게 되면 가중치에 반영될 수 있을 것으로 보았기 때문이다. <표 11>에서 보는 바와 같이 전화와 모바일 조사의 결과가 상당히 근접하는 것으로 나타났다. 즉 본 연구에서 가정했듯이 다양한 측면을 포괄할 수 있도록 가중치를 구할 경우 어느 정도 표본의 편향을 감소시킬 수 있다는 것이다.



<표 11> 가중치 계산 방식에 따른 보정효과

질 문	응 답	전 화	모바일(보정결과)		
			성, 연령, 지역	5개 변인	모든 변인
전자 확인, 미아찾기, 암이나 심장 질환 등과 같은 유전자 정보 관련 문제에 대해 관심이 있으십니까?	관심이 있다	58.1%	70.5%	68.9%	62.8%
	관심이 없다	41.9	29.5	31.1	37.2
만약 귀하가 원하는 정보를 알 수 있다면 귀하는 유전자 검사 를 받을 의향이 있습니까?	받겠다	65.2	82.8	85.1	70.7
	받을 생각이 없다	25.6	14.4	12.0	27.1
	잘 모르겠다	9.2	2.8	2.9	2.2
전반적으로 유전자 검사가 사 회에 미치는 영향은 어떤 것 이 라고 생각하십니까?	불이익이 크다	12.3	21.2	24.2	15.1
	이익이 크다	59.9	68.7	63.7	59.0
	잘 모르겠다	27.8	10.1	12.1	25.9

## V. 결론

이제까지 사용되어 왔던 주류(main) 여론조사방법인 확률적 표본을 사용한 유선전화조사는 이런 저런 문제점을 안고 있어 새로운 대안 모색이 필요한 시점이다. 대안 조사방법으로 휴대전화를 사용한 모바일 패널조사 방법이 제시되고 있다. 이 조사기법은 비확률적 표본 추정기법으로 나름대로의 문제점은 갖고 있지만 노하우(know-how)가 축적되면 결과 예측에 있어 확률적 표본을 사용하는 경우보다 더 정확할 수가 있다. 그 예로 2000년 미국 대선에서 인터넷 패널을 사용한 Harris Interactive의 예측이 유선전화를 사용한 예측결과보다 정확도가 높았던 사실을 들 수 있다.

비확률적 표본을 사용하는 대안조사방법은 왜 정확한 예측결과가 나왔는가를 정교한 표본이론이나 통계이론으로 정확히 설명하지는 못한다. 그러나 많은 의학적 치료기법이 왜 병을 낫게 하는가에 대한 이론적 설명 없이 단지 임상사례(clinical practice)의 결과만 가지고 시술되는 것과 같이 휴대전화 조사나 인터넷 조사도 같은 맥락에서 실제적으로 정확한 예측

이 가능하다면 단지 이론적 배경이 미약하다고 무조건 배척할 수 없다. 의학연구에서는 우선 임상 결과를 다량 축적해 치료에 대한 노하우를 얻어 치료의 효율성을 증가시키고, 그 다음으로 왜 이런 결과가 나왔는가에 대한 이론적 규명을 시도한다. 모바일조사 방법도 아직은 임상사례를 많이 축적하는 단계이다. 앞에서 설명한 대로 이러한 기법은 비확률적 표집들에 따른 전집의 대표성 문제를 수집한 데이터에 의존한 경험적 가중치, 즉 성향지수를 계산해 수정해 준다. 따라서 가중치 계산에 대한 경험적 노하우 축적이 우선 선결되어야 할 문제이다.

이 연구에서 분석한 결과에 따르면 새롭게 등장한 모바일 조사의 표본을 인구주택센서스 자료와 비교해 다양한 측면에서 편향이 나타났다. 즉 현재 모바일 패널 조사에서 사용하고 있는 성, 연령, 거주지역에 입각한 할당표집으로는 대표성 있는 표본을 추출해 주지 못한다는 것이다. 서울과 경기 두 지역의 단체장 선거예측의 경우에는 비교적 오차크기가 적게 나타났지만 다른 지역에서도 그렇게 적게 나타날 것으로 기대하기는 어려웠다. 따라서 모바일조사가 새로운 사회조사 방법으로 자리잡기 위해서는 표집방법론의 개선이 요구된다고 본다. 선거예측의 경우는 할당으로 사용되는 연령과 지역이 투표행위와 밀접한 관련이 있기 때문에 비교적 오차가 적게 나타나지만 다른 사회조사에서는 오차가 더 크게 나타날 수 있기 때문이다.

이러한 맥락에서 흔히 사용되는 가중치 부여방법을 검토하였다. 그 결과 직업과 학력 두 변인만을 사용할 경우에는 효과가 없거나 더 나빠지는 경우도 나타났다. 그러나 12개 변인을 동시에 고려해서 가중치를 산출한 경우에는 전화와 모바일 두 조사의 결과가 근접하였다. 즉 조사오차가 감소하였다. 테일러 등이 주장했듯이 비확률표본에 가중치를 사용함으로써 대표성을 부여하는 것이 가능할 수도 있다는 것을 보여준다.

본 연구의 결과를 볼 때, 보다 다양한 인구학적 속성 및 문화적 변인들을 중심으로 성향지수를 개발하는 노력이 필요하다. 적절한 변인들을 선택할 경우 보다 적은 수의 변인으로 표본의 편향을 제거할 수 있을 것으로 보이기 때문이다.

## 참고문헌

- 강남준. 2002. “전화조사 방법의 대안적 조사기법으로 핸드폰 모바일 조사 : 방법론적 논의 및 실제 사례분석.” 한국조사연구학회·엠비론. 《모바일 서베이 학술세미나 발표논문》.
- 김영원·변종석. “인터넷 조사에서 표본추출 동향 및 문제점.” 한국조사학회. 《인터넷 서베이 워크숍 논문집》.
- 김준한. 2000. “동일 주제에 대한 인터넷 조사와 기존 조사 결과 비교.” 한국조사연구학회. 《인터넷 서베이 워크숍 논문집》.
- 류제복. 2000. “선거예측조사에서 신뢰성 증진방안-16대 총선을 중심으로.” 《조사연구》 1(2) : 35-58.
- 박무익. 2002. “선거여론조사의 역사와 의의.” 《조사연구》 3(1) : 91-118.
- 박용치. 2000. “선거에서 편의의 감소 : 거짓응답을 중심으로.” 《조사연구》 1(2) : 15-34.
- 양승목·김현주·조성겸. 1991. “전화조사에서 가구내 응답자의 무선적 확률표집에 관한 연구 : CNU 선정표의 타당성과 실용성의 검토.” 《신문학보》 26 : 189-214.
- 윤은성·김영원. 2002. “인터넷을 활용한 표본조사 방법에 관한 사례연구.” 《조사연구》 3(1) : 25-51.
- 이계오. 2000. “인터넷 여론조사의 현황과 전망.” 한국조사학회. 《인터넷 서베이 워크숍 논문집》.
- 이준웅. 2001. “여론환경에 대한 인식과 정치적 의견표명 : 제16대 총선 예측조사 실패에 대한 정치커뮤니케이션적 논의.” 《한국방송학보》 15(1) : 199-236.
- 조성겸. 1990. “전화조사 응답자의 표집방법으로서 CNU 방법과 비례확률 표집방법의 비교연구.” 《신문연구소 학보》 27 : 93-106.
- 조성겸. 1997. “선거여론조사의 표본편파 감소방안 : Politz-Simmons 방법과 PSQ 방법의 타당성 검토.” 《언론연구》 7 : 71-98.
- 조성겸. 1997. “대통령 선거 여론조사와 할당표집 방법의 문제점.” 《언론과 사회》 18: 29-53.

- 조성겸. 2001. "2천년 총선 예측조사와 편향요인." 《한국언론학보》 45(3) : 328-359.
- 조성겸. 2001. "선거예측보도의 발전방안." 《방송문화연구》 13 : 281-299.
- 홍내리·허명희. 2001. "제16대 국회의원 선거의 예측조사에 대한 사후적 검증." 《조사연구》 2(1) : 1-36.
- Best, S. F., and B. Krueger. "New Approaches to Assessing Opinion: The Prospects for Electronic Mail Surveys." *International Journal of Public Opinion Research* 14(1) : 73-92.
- Boruch, Robert F., and G. Terhanian. "So What? The Implications of New Analytic Methods for Designing NCEs Surveys." in Gary Hoachlander, Jeanne Griffith and J. Ralph(eds.), *From data to Information : New Directing for the National Center for Education Statistics* 4. Washington, DC : U.S. Department of Education, National Center for Education Statistics.
- Broughton, David. 1995. *Public Opinion Polling and Politics in Britain*. Harvester Wheatsheaf.
- Couper, Mick P. "Web Surveys : A Review of Issues and Approaches." *Public Opinion Quarterly* 64 : 464-494.
- D'Agostino, R. & Rubin, D. (2000) "Estimating and using propensity scores with partially missing data," *Journal of the American Statistical Association* 95 : 749-759.
- Dillman, D. A. 1998. "Mail and Other Self-Administered Surveys in the 21st Century : The Beginning of a New Era." Unpublished Paper. <<<http://survey.sesrc.wsu.edu/dillman/papers.htm>>>
- Frankel, Martin R., and Lester R. Frankel. 1987. "Fifty Years of Survey Sampling in the United Survey." *Public Opinion Quarterly* 51(2) : 127-138.
- Krotki K. & Dennis, M. (2001) "Probability-based Survey Research on the Internet," Paper presented at the 53rd Conference of the International Statistical Institute, Seoul Korea, August : 22-29

- Marsh, Catherine., and Elinor Scarbrough. 1990. "Testing Nine Hypotheses about Quota Sampling." *Journal of the Market Research Society* 32(4): 485-506.
- Mitofsky, Warren. 1999. "Back to the Future of Online Polling: Miscalls Likely in 2000." *Public Perspective* August/September: 42-43.
- Rosenbaum, Paul R., and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Casual Effect." *Biometrika* 70(1) : 41-55.
- Rosenbaum, Paul R., and Donald B. Rubin. 1984. "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score." *Journal of the American Statistical Association* 79(387) : 516-524.
- Rubin, R.(1997). "Estimating causal effects from large data sets using propensity scores." *Annals of Internal Medicine* 127 : 757-763.
- Taylor, H., J. Bremer, C. Overmeyer., J.W. Siegel., and G. Terhanian. 2001. "Using Internet Polling to Forecast the 2000 Elections." *Marketing Research* 13(1) : 26-30.
- Taylor, H. 2000. "Does Internet Research Works," *International Journal of Market Research* 42(1) : 51-63.
- Taylor H., and G. Terhanian. 1999. "Back to the Future of Online Polling: No Witchcraft Here." *Public Perspective* August/September : 42-43.
- Vehovar, Vasja., and K. L. Manfreda. 1999. "Web Surveys : Can the Weighting Solve the Problem?" *1999 Proceedings of the Section on Survey Research Methods*. Alexandria, VA: American Statistical Association.