

연구논문

# 가중방법으로 선거여론조사의 편향을 얼마나 줄일 수 있나: 컴퓨터 시뮬레이션 사례\*

How Much Can Weighting Methods Reduce Bias in Presidential Election Surveys: Results from a Computer Simulation

장덕현<sup>a)</sup> · 홍영택<sup>b)</sup> · 조성겸<sup>c)</sup>

Deok Hyun Jang · Young Taek Hong · Sung Kyum Cho

이 연구는 선거여론조사의 조사결과가 가진 편향을 가중치를 적용하여 감소시키려 할 때, 그 감소폭이 어느 정도인지 그리고 그 감소폭과 가중변인의 설명력 간의 관계는 어떤지를 알아보았다. 이를 위해 가중변인의 설명력을 여러 수준으로 변화시킨 데이터를 컴퓨터로 생성해서 이를 모집단으로 하였다. 그런 다음 관심변인이 10% 포인트의 편향을 갖도록 표본을 반복해서 추출한 다음, 가중치를 적용했을 때 그 편향이 어느 정도 감소하는지를 산출하였다. 그 결과 비록 가중치를 적용하는 것이 편향을 감소시키는 효과가 있기는 하지만 그 효과가 크지 않은 것으로 나타나, 현실적으로 가중치를 적용해서 편향을 제거하는 것은 어려운 것으로 나타났다.

**주제어:** 컴퓨터 시뮬레이션, 가중치, 선거조사

\* 이 논문은 2013년 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF - 2013S1A3A2054988).

- a) 한국갤럽 기획조사실 부장
- b) 한국갤럽 연구6본부 선임연구원
- c) 교신저자(corresponding author): 충남대학교 언론정보학과 교수  
E-mail: skcho@cnu.ac.kr

This study aims to find out the relationship between explanatory power of weighting variables and dependent variable in the context of election poll. We created a population data using previous survey data in which we created the various weighting variables with different level of explanatory power. From this population data, we selected samples repeatedly which have a 10% point bias. When we applied various weighting variables to this sample, we could get the results which have smaller bias. But the reduction is not so big enough for us to get accurate predictions in actual election.

**Key words:** computer simulation, weighting, election poll

## I. 서론

선거예측에서 전화여론조사의 편향을 어떻게 줄일 것인가? 조사의 정확성은 ‘측정치’의 산포를 의미하는 정밀성과 ‘반복된 측정치의 평균값과 참값의 차이’인 편향으로 구성되기 때문에 이들을 분리해서 접근할 필요가 있다 (장덕현·조성겸 2013). 정밀성을 기준으로 할 때 오차를 줄이는 손쉬운 방법은 표본크기를 늘리는 것이다. 표본크기가 커지면 추정치의 분산이 줄고 더 정밀한 측정을 할 수 있다. 혹은 같은 표본크기라 하더라도 적절한 층화를 통해 분산을 줄이는 방법을 고려할 수 있다. 정밀성은 또한 동일한 설계로 이뤄진 조사결과의 일관성을 의미하기 때문에 조사의 과정에서 일관된 방법과 절차를 적용하는 것이 효과적일 수 있다.

그러나 측정치의 평균과 참값의 차이인 편향에 대해서는 표본크기는 관련이 없고, 조사과정의 표준화도 편향의 여러 요인 중의 하나일 뿐이다. 따라서 표본크기를 키우더라도 편향은 줄어들지 않게 되고, 조사과정의 표준화 역시 편향을 감소시키는 충분조건은 되지 않는다.

편향을 줄이는 가장 유력한 방법은 확률표집 등 과학적 원칙을 충실히 준수하는 것이고 현재의 조사환경에서 가장 중요한 요소는 응답률이다. 한국의 선거여론조사에서는 대부분 응답률을 높이기 위한 재통화(callback)보다는 할당을 설정하고 대체하는 방식을 취한다(박민규 외 2012; 허명희 외 2011). 할당이라는 제한이 있지만 이처럼 무한정 대체를 한다면 응답 협조도가 높은 특성을 가진 표본이 과도하게 표집될 가능성이 높다. 그렇다고 해서 재통화를 실시하는 것이 대안이 되는 것은 아니다. 재통화를 통해 비수신 전화를 포함시키고, 거절자를 설득한다면 편향이 줄어들 가능성이 있지만 이러한 방식으로 응답률을 높이는 것이 조사의 편향을 줄인다는 경험적 증거는 실제로 존재하지 않는다. 물론 응답률을 80%나 90% 수준으로 높이면 좋겠지만 현실적으로는 이렇게 높은 응답률을 얻는 것이 가능하지 않기 때문이다.

다른 유력한 방법은 적절한 성향 가중을 통해 보정하는 것이다. 성향 가중은 응답 여부(응답확률,  $P$ )에 영향을 주는 공변량( $Z$ )을 통해 조사변수( $Y$ )의 편향을 줄이려는 접근이다(Rosenbaum & Rubin 1984; Lee 2006). 만약 만족할 만한 정밀성을 확보했다면 성향 가중을 통해 편향을 줄여 정확한 예측을 할 수 있을 것이다. 실제로 자원자 패널을 이용한 인터넷 조사결과에 성향 가중치를 적용해서 선거결과를 예측하는 데 활용한 사례가 국내에도 있다(Traugott 2001; 김원용·이홍철 2003; 이계오·장덕현 2009). 그렇지만 이러한 시도에도 불구하고 이러한 방법이 예측에 도움을 준다는 명확한 증거는 아직 없다.

성향 가중방법에서 성향점수는 조사에 응답한 확률에 대한 추정치를 의미한다. 따라서 성향가중을 하면 응답확률이 높은 표본의 비중을 줄이고 응답확률이 낮은 표본의 비중을 늘려 모집단 특성에 근접할 수 있다. 그러나 응답확률을 통해 보정하더라도 성향 가중을 통해 효과적으로 편향을 줄이기 위해서는 조사변수에 충분히 영향을 주는 성향점수여야 한다. 조사변수와 연관성이 없다면 성향 가중을 통한 보정효과를 기대하기 어렵다.

그렇다면 성향점수와 조사변수의 연관성은 어느 정도여야 하는가? 예측의 정확성을 높이기 위해 상관성이 어느 정도로 높은 성향변수를 찾아야 하

는가? 본 연구에서는 이를 확인하기 위해 컴퓨터 시뮬레이션을 시도했다. 구체적으로 가상의 모집단을 구성하고 일정한 편향을 포함한 표본을 추출했으며, 목표변수와 일정한 상관성을 가진 성향변수를 통해 추출된 표본을 보정하는 과정을 거쳤다. 상관성에 따라 보정효과가 얼마나 되는지, 즉 그 크기가 통계적으로뿐만 아니라 실제적으로 유의미한지 확인하고자 한다.

## II. 이론적 배경 및 기존연구 검토

### 1. 무응답률과 편향과의 관계

전화여론조사에서 다양한 요인들이 편향을 발생시킬 수 있다. 예를 들어, 질문의 의미가 모호하여 응답자가 체계적으로 잘못 이해하거나 사회적 환경의 영향으로 편향된 응답을 할 수 있다. 또한 전화번호 등재율이 낮은 전화번호부를 사용하는 조사와 같이 표본추출틀의 포함률에 문제가 있을 수 있다.

무엇보다도 전화조사에서 표본의 편향 발생원인으로는 비수신이나 거절, 중단을 포함한 무응답률(non-response)이 의심된다. 그 이유는 전화조사의 무응답률이 매우 높기 때문이다. 휴대전화 RDD 방식을 기본으로 진행되는 한국갤럽의 데일리 조사에서 2013년 1월 이후 주간 응답률(COOP)<sup>1)</sup>은 20%를 넘지 못한다. 무응답률이 높기 때문에 응답자와 무응답자 간에 체계적인 차이가 있다면 결과가 크게 달라질 수 있다. 반면 다른 이유에서는 편향이 발생할 가능성은 크지 않다. 조사질문 등 진행상의 문제가 지속적으로 발생할 가능성은 적고, RDD 방식의 적용으로 전화번호부를 기반으로 하던 기존 전화조사의 모집단 포함률 문제는 상당 부분 극복됐다.

그렇다면 무응답률과 조사의 오차 및 편향과의 관계는 어떠한가? 전화조사의 응답률을 높이는 것이 편향을 얼마나 줄일 수 있을까? 응답률이 높다는 것은 무응답에 의해 편향이 발생할 가능성이 낮다는 것을 의미한다. 그

1) AAPOR 응답률 기준  $COOP1 = I / \{(I + P) + R + O\}$

러나 결론적으로 말하면, 응답률을 다소 높인다고 해서 편향을 반드시 제거할 수 있는 것은 아니다. 다수의 연구들에서 응답률과 응답편향과의 관계를 발견할 수 없었다(Curtin et al. 2000; Keeter et al. 2000; Merkle & Edleman 2002). Keeter et al.(2000)의 연구에서 표준 방식으로 5번 재통화하여 응답률 36%를 달성한 방식과 엄격하게 8주 동안 60.6% 응답률을 달성한 방식을 비교했다. 그 결과 전체 91개 항목 중 15개 항목에서 통계적으로 유의한 차이가 있었지만 실제 차이는 대부분 5%p 미만으로 차이가 크지 않았다. Groves & Peytcheva(2008)가 59개 연구, 959개 무응답편향 추정치를 메타 분석한 결과에서도 응답률과 편향과의 관련성을 발견할 수 없었고, 결론적으로 응답률의 크기 그 자체가 아니라 조사 주제에 대한 응답자의 관심 등 무응답이 발생한 원인이 중요하다고 밝히고 있다.

밀워드브라운미디어리서치가 2013년 4월 휴대전화 RDD 방식으로 최대 11회까지 재통화한 결과에서도 첫 번째 통화의 응답자와 최종 응답자의 특성별 분포에서 의미있는 차이를 발견할 수 없었다(Cho & Jang 2013).

이것은 재접촉 시도 등을 하더라도, 여전히 응답성향이 높은 사람이 응답을 더 하게 되기 때문이다. 무응답이 편향을 초래하는 것은 응답한 사람과 응답하지 않은 사람 간의 성향 차이가 있기 때문인데, 재접촉 시도에서도 여전히 처음 접촉할 때와 마찬가지로 편향이 나타나기 때문이다. 즉 재접촉 등으로 응답률을 현저히 높이지 않는 한 응답률과 편향은 관련성이 없게 된다. 만약 재통화를 통해 응답률을 높이는 것이 편향을 줄이는 데 효과적이지 않다면 남은 방법은 무응답편향이 발생하는 원인을 찾아 성향 가중을 통해 보정하는 것이다.

## 2. 무응답편향과 성향 가중

무응답편향에 대해서는 3가지 모형이 제시된 바 있다(Groves & Peytcheva 2008). 첫째는, 분리 원인 모형(seperate cause model)으로 조사변수인  $Y$ 의 발생 원인과 응답성향인  $P$ 의 발생 원인이 독립적임을 의미한다. 이런 경우 응답자로부터 얻어진  $Y$ 의 기대치는 편향되지 않고 응답자로부터 측정된 것

이나 무응답자로부터 측정된 것이나 차이 나지 않는다.

둘째, 공통 원인 모형(common cause model)으로  $Y$ 와  $P$ 가 공통적인 원인( $Z$ )을 가지고 있으므로 공통 원인을 통해  $Y$ 를 보정할 수 있다. 예를 들어, 60세 이상의 무응답률이 높고 이들이 보수 후보에게 투표하는 성향이 있다면 연령대는 공통 원인이 된다. 이때는 60세 이상의 비중을 늘리면  $Y$ 를 보정할 수 있다.

셋째, 조사변수 원인 모형(survey variable cause model)으로 조사변수  $Y$ 가 무응답( $P$ )의 원인이 되는 경우다. 예를 들어, 무응답의 원인이 되는 성, 연령, 지역 등 다른 변수와 상관없이 특정 후보에 투표할 사람들이 응답하지 않는 경우다. 이때는 성향을 통해  $P$ 를 보정하더라도 조사변수  $Y$ 는 변하지 않는다.

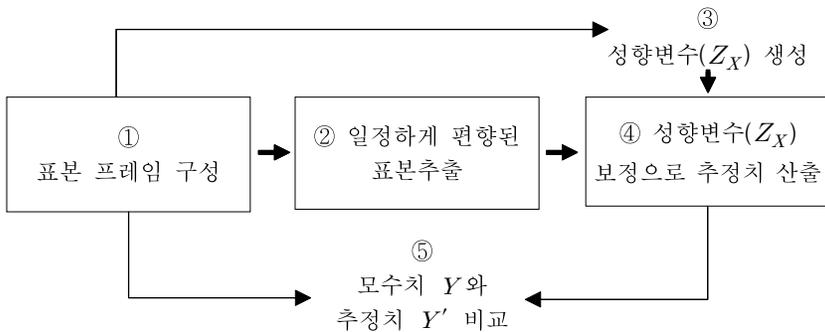
성향점수 가중 적용은 공통 원인 모형을 전제한다. 공통 원인 모형이 적용되기 위해서는 응답성향뿐 아니라 조사변수인  $Y$ 와 높은 상관성을 보이는 공통 원인 변수를 찾아야 한다.

인터넷조사와 같은 비확률추출방법의 경우 성향점수 가중을 통해 편향을 보정한 다양한 사례들이 존재한다. 대표적으로 해리스인터랙티브는 성향점수 가중을 통한 인터넷조사로 2000년 미국 대선결과를 예측하였다(Traugott 2001). Soest et al.(2007)의 연구에서는 개방성이나 사생활 침해에 대한 생각, 여행, 스포츠, 독서와 같은 활동 등을 이용하였는데, 이는 웹조사의 응답 확률과 관련된 질문들이다. 이계오와 장덕현(2009)은 2007년 대통령 선거에 대한 인터넷 조사에 대해 성향점수 가중을 적용한 바 있다. 이때 성향점수의 도출에 사용했던 변수들은 국가보안법 폐지, 대형주택 중과세, 북미관계 호전, 대북 지원, 한미 FTA 등 사회적 이슈에 대한 태도와 이념성향, 미국, 일본, 중국, 북한 등 주변 국가에 대한 태도 등인데, 모두가 대선 후보 지지도, 지지 정당과 같은 정치적 성향과 관련성이 큰 변수들이었다.

공통 원인 모형을 가정한다면 성향변수는 응답확률과 조사변수 모두에 관련되어야 한다. 그렇다면 어느 정도 관련성이 있어야 편향을 보정할 수 있을까? 현실에서는 다양한 사례를 시도할 수 없기 때문에 현실과 유사한

가상의 상황을 설정해야 한다. 이는 시뮬레이션을 통해 가능하며 본 연구에서는 2012년 대선 후보 지지도 자료를 통해 이를 시도했다.

### Ⅲ. 연구방법



<그림 1> 시뮬레이션 연구 절차

우리는 모집단의 표본 프레임을 구성해 표본추출을 하고 추정치를 산출하는 과정을 컴퓨터 시뮬레이션을 통해 실행하려 한다. 특히 본 연구의 목적이 성향 가중 보정의 효과에 있으므로 일정하게 편향된 표본을 추출할 것이며, 조사변수인  $Y$ 와 일정한 관련성을 가진 성향변수( $Z$ )를 생성할 것이다.

먼저, 표본 프레임은 2012년 10월 15일부터 선거 직전일(D-1)인 12월 18일까지 한국갤럽 데일리 조사 프로그램의 응답자 19,709명으로 구성했다. 이들은 각 시기별로 지지 정당과 대선 후보 지지 성향을 밝혔으며 전화조사의 할당 기준으로 사용하는 지역, 성, 연령대가 파악되어 있다. 또한, 정치적 성향과 관련성이 높은 특성인 직업, 주관적 생활수준, 18대 총선 투표 후보가 조사됐다. 모집단의 전반적인 특성 분포는 <표 1>과 같다.

〈표 1〉 표본 프레임의 특성 분포

(단위: 명)

		사례 수	%			사례 수	%
전체		19,709	100.0	전체		19,709	100.0
지역	서울	4,351	22.1	직업	농/임/어업	672	3.4
	인천/경기	5,753	29.2		자영업	3,340	16.9
	강원	570	2.9		블루칼라	2,185	11.1
	대전/충청	1,868	9.5		화이트칼라	5,797	29.4
	광주/전라	1,990	10.1		가정주부	4,750	24.1
	대구/경북	2,004	10.2		학생	1,308	6.6
	부산/울산/경남	2,984	15.1		무직/기타	1,657	8.4
	제주	189	1.0		상/중상	2,277	11.6
성	남성	10,479	53.2	주관적 생활수준	중	8,374	42.5
	여성	9,230	46.8		중하	4,988	25.3
연령대	19~29세	2,915	14.8		하	3,585	18.2
	30대	3,799	19.3		모름/응답거절	1,657	2.5
	40대	4,410	22.4		총선투표 후보	새누리당	6,853
	50대	4,559	23.1	민주통합당		6,307	32.0
	60세 이상	4,026	20.4	통합진보당		633	3.2
			기타/무소속	651		3.3	
지지 후보	박근혜	9,327	47.3	투표 안함/ 모름/응답거절	5,264	26.7	
	문재인	8,498	43.1				
	기타 후보	110	0.6				
	모름/응답거절	1,774	9.0				
지지 정당	새누리당	7,453	37.8				
	민주통합당	5,872	29.8				
	기타 정당	690	3.5				
	모름/응답거절	5,694	28.9				

둘째, 19,709명 표본프레임에서 1,000명의 표본을 31회 반복 추출했다.<sup>2)</sup> 표본추출방법은 지역/성/연령대 크기 비례 층화확률추출을 적용했다. 단 각 층별로 박근혜 후보 지지자의 추출확률을 증가시킴으로써 표본의 조사변수(박근혜 후보 지지도)의 편향을 10%포인트 가량 발생시켰다.

셋째, 지지 후보와 일정한 관련성을 가진 성향변수( $Z$ )를 생성했다. 성향변수의 생성 절차는 다음과 같았다.

- ① 표본 프레임에서 후보 지지도와 완벽히 일치하는 성향변수  $Z_0$ 를 생성한다.  
if (지지후보=1)  $Z_0=1$ , else=2
- ② 1에서 20까지의 값을 가진 난수( $XX$ )를 생성한다.
- ③ 난수값에 따라  $Z_0$ 의 값을 변화시킨 새로운  $Z_X$  변수를 생성한다.  $Z_X$ 는 모두 7개였는데, 그 중  $Z_1$ 이 가장 높은 설명력을 그리고  $Z_7$ 이 가장 낮은 설명력을 갖도록 하였다. 이와 같은 변수생성 과정을 SPSS의 syntax형식으로 제시하면 다음과 같다.

```

if (XX ≤ 1 and Z0 = 1) Z1 = 2
if (XX ≤ 1 and Z0 = 2) Z1 = 1
if (XX > 1) Z1 = Z0
if (XX ≤ 2 and Z0 = 1) Z2 = 2
if (XX ≤ 2 and Z0 = 2) Z2 = 1
if (XX > 2) Z2 = Z0
...
if (XX ≤ 7 and Z0 = 1) Z7 = 2
if (XX ≤ 7 and Z0 = 2) Z7 = 1
if (XX > 7) Z7 = Z0

```

---

2) 표본은 컴퓨터 시뮬레이션을 통해 추출됐지만 추출된 표본을 토대로 가중치를 산정하고 그것을 적용해서 오차변화를 살펴보는 작업은 별도로 이루어져야 했다. 따라서 일반적인 컴퓨터 시뮬레이션 연구처럼 1,000회 등과 같은 많은 횟수를 반복하는 것은 어려웠다. 또 본 연구에서는 31회의 반복추출로도 필요한 변량을 확보할 수 있다고 판단했다.

〈표 2〉 생성변수( $Z_0$  대  $Z_1 \sim Z_7$ ) 교차표

(단위: 사례 수)

		$Z_1$		$Z_2$		$Z_3$		$Z_4$	
		1	2	1	2	1	2	1	2
$Z_0$	1	8,844 (94.7)	497 (5.3)	8,383 (89.1)	1,022 (10.9)	7,883 (83.9)	1,513 (16.1)	7,431 (78.4)	2,050 (21.6)
	2	483 (4.7)	9,885 (95.3)	944 (9.2)	9,360 (90.8)	1,444 (14.0)	8,869 (86.0)	1,896 (18.5)	8,332 (81.5)
		$Z_5$		$Z_6$		$Z_7$			
		1	2	1	2	1	2		
$Z_0$	1	6,960 (73.1)	2,557 (26.9)	6,494 (67.8)	3,081 (32.2)	6,070 (62.5)	3,641 (37.5)		
	2	2,367 (23.2)	7,825 (76.8)	2,833 (28.0)	7,301 (72.0)	3,257 (32.6)	6,741 (67.4)		

\* ( ) 안은 Row %임

넷째, 각 표본에 대해 성향점수 보정을 통해 조사변수인 지지후보의 추정치를 산출한다. 각 표본은 지역, 성, 연령대는 설계변수로 투입됐으므로 각 셀의 크기를 일치시킨 설계가중 추정치( $Y_D$ )를 기본으로 한다. 이때 박근혜 후보 지지도의 기댓값은 모수치에 비해 10%포인트 가까운 편향을 가진다. 직업, 주관적 생활수준, 총선 투표 후보는 이미 알 수 있는 통제변수로 간주한다. 따라서, 이들 변수를 통제하기 위해 림 가중을 적용한 추정치( $Y_C$ )를 산출한다. 이처럼 가중변수가 여러 개일 경우 하위 교차 셀의 빈도가 0이 나타날 수 있어 셀 가중치를 산출할 수 없다. 따라서 각 변수의 한계분포만을 일치시키는 림 가중치를 적용한 것이다(Lohr 1999). 다시 새롭게 생성한 성향변수인  $Z_1$ 부터  $Z_7$ 까지 포함해 림 가중을 적용하여 성향변수 가중이 추가된 추정치( $Y_{z_1} \sim Y_{z_7}$ )를 산출한다. 결과적으로 31개 각 표본에 대해 설계가중 추정치( $Y_D$ ), 통제변수 가중 추정치( $Y_C$ )와  $Y_{z_1}$ 부터  $Y_{z_7}$ 까지 7개의 성향가중 추정치 등 모두 합쳐 9개의 추정치를 산출할 것이다.

마지막으로, 편향된 표본 추정치( $Y_D$ )로부터 성향가중을 통해 편향이 얼마나 줄었는지를 확인할 것이다. 여기서 편향은 표본 추정치 평균의 오차로

규정한다. 우리가 생성한 성향변수 중  $Z_1$ 은 가장 강력한 변수이다. 문제는 이렇게 조사변수와 높은 상관성을 가진 변수를 실제로 찾을 수 있는가이다. 시뮬레이션 결과 분석은 현실적으로  $Y_{z_1}$ 부터  $Y_{z_7}$ 에서 달성할 수 있는 목표 수준을 설정하게 해 줄 것이다.

#### IV. 결과 분석

##### 1. 조사변수와 성향변수의 관계

로지스틱 회귀분석을 통해 조사변수에 대해 생성한 성향변수가 얼마나 관련됐는지를 확인했다. 조사변수와 성향변수의 관련성에 따라 성향 가중치 편향을 줄이는 데 효과적인지를 파악하게 될 것이다.

<표 3> 로지스틱 회귀분석 -  $R^2$ , 우도비 검정 결과

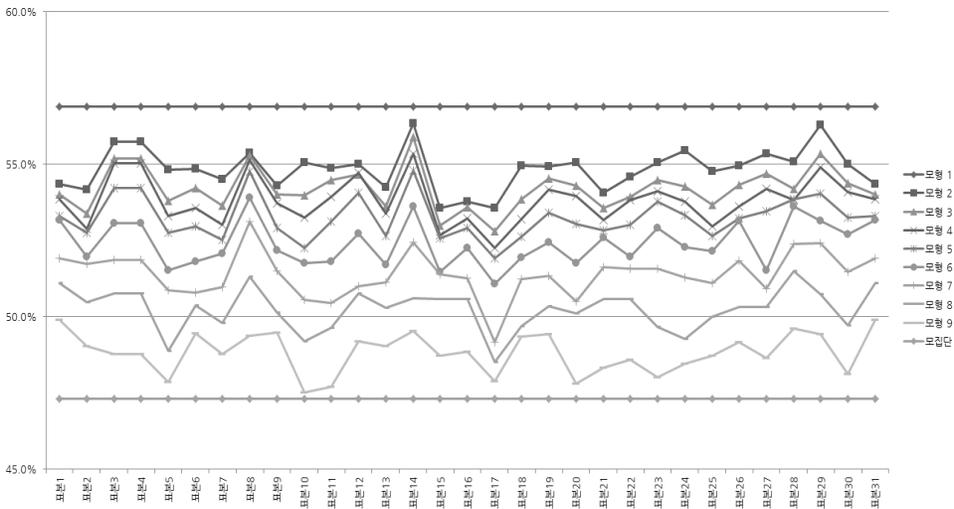
		모형1	모형2	모형3	모형4	모형5	모형6	모형7	모형8	모형9
		설계 변수	통제 변수	$Z_7$	$Z_6$	$Z_5$	$Z_4$	$Z_3$	$Z_2$	$Z_1$
Pseudo $R^2$	Cox & Snell	.172	.347	.384	.412	.447	.490	.538	.596	.661
	Nagelkerke	.229	.464	.513	.549	.596	.653	.718	.795	.881
	McFadden	.136	.309	.351	.383	.428	.486	.558	.655	.780
우도비 검정 ( $p$ )	지역	.000	.000	.000	.000	.000	.000	.000	.000	.000
	성/연령대	.000	.000	.000	.000	.000	.000	.000	.000	.000
	생활수준		.000	.000	.000	.000	.001	.003	.013	.049
	직업		.000	.000	.000	.000	.000	.001	.005	.046
	총선투표 후보		.000	.000	.000	.000	.000	.000	.000	.000
	성향변수 ( $Z_7 \sim Z_1$ )			.000	.000	.000	.000	.000	.000	.000
박근혜 후보 분류 정확률(%)		66.7	74.4	75.7	78.5	81.7	84.4	86.9	89.9	94.8

먼저 설계변수인 지역과 성/연령대가 조사변수를 얼마나 설명하는지를 알아보고, 직업, 주관적 생활수준, 총선 투표 후보와 같은 통제변수를 추가했을 때 모형의 설명력이 어떻게 변하는지 볼 것이다. 그 다음  $Z_7$ 부터  $Z_1$ 까지 성향변수를 각각 추가하여 모형의 설명력을 확인할 것이다. 종속변수는  $Z_0$ (지지후보=1일 때 1, 이 외는 2)로 설정했다.

생성된 성향변수는 우도비 검정 결과 모두 유의했으며 통제변수가 추가된 모형2 대비 모형의 설명력을 증대시켰다. Nagelkerke  $R^2$  기준으로 모형의 설명력을 살펴보면 통제변수를 추가한 모형2는 .464,  $Z_7$ 을 추가한 모형3은 .513로 .05 가량 설명력이 증가했으며  $Z_1$ 을 추가한 모형9는 최대 .881까지 증가했다.

## 2. 성향 가중의 결과

1,000명씩 31차례 표본을 반복 추출했으며 각 표본은 9.6% 포인트 박근혜 후보 지지자가 과다 표집되도록 했다. 따라서 모집단의 박근혜 후보 지지도



〈그림 2〉 표본별 림 가중의 편향제거 효과(박근혜 후보 지지도 추정치)

&lt;표 4&gt; 림 가중의 편향 제거 효과(편향의 크기)

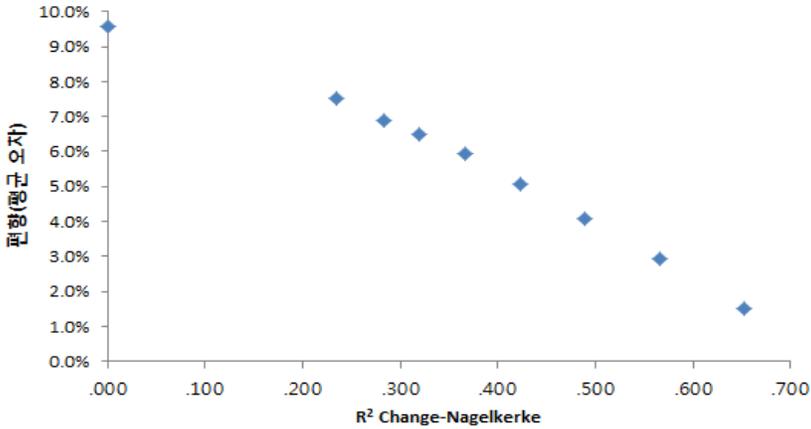
	모형1	모형2	모형3	모형4	모형5	모형6	모형7	모형8	모형9
	설계 변수	통제 변수	$Z_7$ 추가	$Z_6$ 추가	$Z_5$ 추가	$Z_4$ 추가	$Z_3$ 추가	$Z_2$ 추가	$Z_1$ 추가
박근혜 후보 지지도 추정치 표본평균	56.9	54.8	54.2	53.8	53.2	52.4	51.4	50.3	48.8
표본 표준편차	0.0	0.7	0.7	0.8	0.7	0.7	0.7	0.7	0.7
최댓값	56.9	56.3	55.9	55.4	54.8	53.9	53.1	51.5	49.9
최솟값	56.9	53.6	52.8	52.3	51.9	51.1	49.2	48.5	47.5
편향(평균 오차)	9.6	7.5	6.9	6.5	5.9	5.1	4.1	2.9	1.5

는 47.3%이고 설계가중에 의한 박근혜 후보 지지도 추정치( $Y_D$ )는 56.9%로 모든 표본에 대해 동일하다.

각 표본별로 직업, 주관적 생활수준, 총선 투표 후보의 통제변수와  $Z_1$ 부터  $Z_7$ 까지 성향변수 분포를 통해 림 가중한 결과는 <그림 2>와 같다. 조사 변수와 관련성이 높은 성향변수를 추가할수록, 즉 모형의 설명력이 증가할수록 지지도 추정치는 모집단 추정치에 가까워졌다.

통제변수를 통해 림 가중했을 때 추정치( $Y_C$ )의 평균(모형2)은 54.8%로 편향(평균 오차)은 2.1% 포인트 줄어 7.5% 포인트였다.  $Z_7$  성향변수를 추가해 림 가중했을 때(모형3) 추정치( $Y_{z_7}$ ) 평균은 54.2%로 모형2와 차이가 거의 없었으나,  $Z_4$ 를 추가한 모형6에서 편향이 5.1% 포인트로 줄고  $Z_1$ 를 추가한 모형9에서 추정치( $Y_{z_1}$ )의 편향이 1.5% 포인트까지 줄었다.

모형의 설명력과 대비해 보면 <그림 3>과 같이 설명력( $R^2$ )이 증가할수록 편향이 선형적으로 감소함을 볼 수 있다. 추정치의 편향이 절반 가까이 감소한 모형6의 경우  $R^2$ 는 .653에 달하고 분류 적중률은 84.4%에 이른다(<표 3>참고). 즉, 가중방법으로 편향을 절반 가까이 줄이기 위해서는  $Z_4$  정도로 높은 설명력을 가진 성향변수가 필요하다는 것을 말한다.



〈그림 3〉 모형 설명력 변화에 따른 표본 추정치의 편향

우리는 모형의 설명력을 증가시키는 가상의 성향변수를 생성하여 가중처리했을 때 추정치의 편향이 얼마나 줄어드는지 시뮬레이션을 통해 확인해 봤다.

시뮬레이션 결과는 가중방법을 통한 보정이 편향을 줄이는 데 그다지 효과적이지 않다는 것을 보여준다. 즉, 편향이 일단 발생하면 조사변수에 대해 상당히 높은 설명력을 가진 성향모형이 필요한데 현재 현실적으로 설정 가능한 통제변수 가중 하에서 단지 2.1% 포인트의 편향(전체 편향의 21.9%)을 줄였을 뿐이다. 또한, 추정치의 편향을 절반 가까이 줄이기 위해서  $R^2$ 를 .65 이상 달성해야 하는데 뒤에 검토하겠지만 이는 현실적으로 얻기 힘들다.

## V. 결론

조사과정에서 발생한 편향은 가중치를 적용해서 감소시킬 수 있는 것으로 간주되어 왔다. 그렇기 때문에 가중치를 적용하면 과연 편향이 실제 감

소되는지 여부에 대해서는, 특히 선거여론조사 상황에서 체계적으로 검토되지 않았다.

본 연구는 모바일 조사를 실시할 때 나타나는 무응답으로 인한 편향을 가중치를 통해 제거하기 위해서는 어떤 변인들을 가중치로 사용해야 하는가를 체계적으로 검토해 보았다. 그 결과 일반적으로 널리 사용되는 지역, 성/연령대만을 통제했을 경우에는 체계적 편향을 거의 감소시키지 못하는 것으로 나타났고, 직업, 생활수준 및 과거 투표성향과 같이 현재 측정이 가능한 여러 변수를 통제했을 경우에 겨우 25% 포인트의 편향을 감소시키는 것으로 나타났다. 이런 정도라면 두 후보 간 격차가 대개 4~5% 포인트 정도인 선거구의 경우 예측치로서 도움이 되지 않는다. 소위 경합지역으로 분류되는 그래서 두 후보 간의 격차가 5% 포인트 정도일 것으로 예상되는 지역에서 활용되기 위해서는 50% 이상 오차를 감소시킬 수 있어야 하는데 그러기 위해서는 Pseudo  $R^2$ 이 .60 이상의 설명력을 가져야 한다. 그러나 현재 조사에서 주로 측정되는 과거 투표성향, 직업, 생활수준 등을 포괄적으로 적용한다고 해도 이러한 설명력은 얻기 어렵다. 이렇게 볼 때 현재 정치여론조사에서 측정되는 변인들을 토대로 가중치를 적용하는 것은 그다지 편향 감소의 의미가 높다고 볼 수 없다.

지금까지 선거과정에서 발생하는 무응답 그리고 이로 인한 편향은 가중치를 적용해 제거할 수 있다고 믿어왔다. 그러나 본 연구의 결과는 가중치를 그렇게 신뢰하기 어렵다는 것을 보여준다. 물론 무응답률을 획기적으로 낮추면 좋겠지만 그것은 현실적으로 어려운 문제다. 또한 편향이 개입되는 것은 무응답의 크기 그 자체보다는 무응답의 분포가 얼마나 편향되었는가에 달렸다. 따라서 무응답 자체를 줄이려는 노력도 중요하지만 무응답 분포에 체계적 차이를 유발하는 변인들을 찾아내고 조사과정에서 그러한 변인의 영향을 통제하는 것이 조사의 정확성 제고에 중요하다.

또한 조사결과의 편향 여부 역시 무응답 분포를 통해 어느 정도 추정할 수 있지만, 현재 일부 조사를 제외한 대부분은 할당표집을 하기 때문에 무응답 분포를 파악하기 어렵다. 따라서 앞으로 조사과정에서 이러한 무응답 분포를 파악하는 노력이 더욱 중요하다고 할 수 있다.

## 참고문헌

- 김원용 · 이홍철. 2003. “웹조사의 모집단대표성 확보를 위한 성향가중 모형의 적합성 검증.” 《방송연구》 여름호: 143-166.
- 박민규 · 조성겸 · 송종현 · 김옥태 · 장운재. 2012. “할당추출표본을 이용한 추정량의 모형 편향에 대한 연구.” 《조사연구》 13(2): 99-109.
- 이계오 · 장덕현. 2009. “인터넷 선거조사에서 성향가중모형 적용사례.” 《조사연구》 10(3): 21-36.
- 장덕현 · 조성겸. 2013. “무선 RDD 전화조사의 타당성 검토 연구: 18대 대통령 선거 사례.” 《조사연구》 14(3): 19-47.
- 허명희 · 한상태 · 김지연 · 성은하 · 강현철. 2011. “임의번호걸기와 시간균형할당 표집에 의한 전화조사의 주요 결과.” 《조사연구》 12(2): 77-88.
- Cho, S. K. and D. H. Jang. 2013. “Opinion Surveys Conducted Using Mobile Phones with Supplementary Landline Survey.” KOSTAT Conference.
- Curtin, R., S. Presser, and E. Singer. 2000. “The Effect of Response Rate Changes on the Index of Consumer Sentiment.” *Public Opinion Quarterly* 64: 413-428.
- Keeter, S., C. Miller, A. Kohut, R. M. Groves, and S. Presser. 2000. “Consequences of Reducing Nonresponse in a National Telephone Survey.” *Public Opinion Quarterly* 64. 125-148.
- Merkle, D. and M. Edelman. 2002. “Nonresponse in Exit Polls: A Comprehensive Analysis.” In R. M. Groves, D. A. Dillman, J. L. Eltinge, and R. J. A. Little(Eds). *Survey Nonresponse*. pp. 243-258. New York: Wiley.
- Groves, R. M. and E. Peytcheva. 2008. “The Impact of Nonresponse Rates on Nonresponse Bias: Meta Analysis.” *Public Opinion Quarterly* 72(2): 167-189.

- Lee, S. 2006. "Propensity Score Adjustment as a Weighting Scheme for Volunteer Panel Web Surveys." *Journal of Official Statistics* 22(2): 329-340.
- Lohr. S. H. 1999. *Sampling: Design and Analysis*. Duxbury Press.
- Rosenbaum, P. R. and D. B. Rubin. 1984. "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score." *Journal of the American Statistical Association* 79(387): 516-524
- Soest A. V., M. Schonlau, and A. Kapteyn. 2007. "Are 'Webographic' or Attitudinal Questions Useful for Adjusting Estimates from Web Surveys Using Propensity Scoring?" *Survey Research Methods* 1(3): 155-163.
- Traugott, M. W. 2001. "Assessing Poll Performance the 2000 Campaign." *Public Opinion Quarterly*. 65(30): 389-419.

<접수 2013/11/6, 수정 2014/2/4, 게재확정 2014/2/8>