

연구논문

자원자 패널에 의한 인터넷 조사의 성향조정 가중화

Propensity Adjustment Weighting of the Internet Survey by Volunteer Panel

허명회* · 조성겸**

Myung-Hoe Huh · Sung Kyum Cho

본 연구는 인터넷 자원자 패널에 의한 사회조사 사례를 보고하면서 조사 자원자의 사회심리적 특성을 규명하고 패널표본의 편향성을 제거하기 위한 성향조정 가중법(propensity adjustment weighting)의 효용에 대한 정량적 평가 결과를 제시한다. 광역시·도, 지역크기(동·읍면), 성별, 연령대, 교육정도, 소비생활 만족도, 소득분배 견해, 일반신문 보기, 인터넷신문 보기 등 9개 요인을 조정 기준변수로 사용하여, 2009년 통계청 온라인 사회조사 자료(응답자 수 2,903명)를 같은 해에 지역추출 면대면(face-to-face, FTF)으로 실시된 통계청 사회조사 자료(응답자 수 37,049명)에 맞도록 성향조정 가중치를 적용하였다. 그 결과 106개 평가 문항 가운데 80% 정도의 문항에서 광역시·도, 동·읍면, 성별, 연령대 등 기본 인구지리적 변수에 의한 기본가중치 보정에 비하여 향상된 추정치를 제공하는 것으로 평가되었다. 이는 성향조정에 일부 한계가 있기는 하지만 자원자 패널에 의한 인터넷 조사를 향후 사회조사에서 활용해 볼 가치가 있음을 의미한다. 본 연구에서 준거조사(reference survey)로 쓰인 2009년 통계청 면대면 사회조사는 표본크기가 크고 일반적인 사회심리적 태도 문항과 인터넷 관련 행태 문항을 포함하므로 각종 인터넷 조사에서 준거자료로서의 활용가치가 충분히 있다.

주제어: 인터넷 조사, 사회조사, 자원자 패널, 성향조정 가중법, 로짓모형, rim 가중법

This paper reports the results of the 2009 Internet volunteer panel version of the social survey conducted by Statistics Korea (Korea National Statistical Office). Authors identify socio-psychological characteristics of Internet survey volunteers and present quantitative evaluation of the propensity adjustment weighting method intended to remove Internet sample bias. The nine criteria used for propensity adjustment were regions, urban/rural, gender, age, education, consumer satisfaction, views on income

* 교신저자(corresponding author): 고려대학교 정경대학 통계학과 교수 허명회.

E-mail: stat420@korea.ac.kr

** 충남대학교 사회과학대학 언론정보학과 교수

distribution, newspaper access and Internet news access. Propensity adjustment weighting based on the logit model and rim weights were applied to the online survey of 2,903 respondents using the face-to-face area sample data of 37,049 respondents as reference. A total of 106 items were used for evaluating the propensity adjustment weighting methods. The results showed that in 80% of survey items the propensity adjustment weighting yielded better estimates compared to simple demographic weighting. This suggests that Internet surveys by volunteer panels are useful for conducting the general social study in Korea. The reference survey data for this study contains several items on social-psychological behaviors and attitudes, is large in size and obtained by probability sampling. Thus it may be utilized in propensity adjustment of other Internet surveys.

Key words : Internet survey, social survey, volunteer panel, propensity-adjustment weighting, logit model, rim weighting

I. 연구 배경과 목적

우리나라에서 공식통계 작성을 위한 표본조사는 대부분 가구를 직접 방문하여 면접타계식으로 조사하고 있다. 그러나 1인 가구(20%)와 맞벌이 가구(25%)의 증가로 인해 부재중인 가구가 늘어나고 오피스텔 및 아파트 공동현관에 무인경비시스템을 설치하는 경우도 증가함에 따라 점점 면접원이 접촉하기 어려운 가구가 많아지고 있다. 또한 면접원에 의한 사생활 침해를 우려하는 경우도 늘고 있어, 조사환경이 급속히 악화되고 있는 추세에 있다.

이를 타개하는 한 방법으로서 자원자 패널(volunteer panel)에 의한 인터넷 조사를 고려할 수 있다. 그러나 이런 방식의 인터넷 조사는 크게 두 가지의 오차요인을 갖는다. 첫째는 포함률 오차(coverage error)로서 인터넷에 접근할 수 있는 응답자로 표본이 제한됨으로써 발생하는 오차이다. 둘째는 응답자의 자기선택 오차(self-selection error)로서 표본추출이 응답자의 자발성에 의존함으로써 발생하는 오차이다. 즉 다른 전통적인 조사방법보다는 응답자의 협조여부가 표본선정 과정에서 중요한 역할을 하게 된다는 것이다. 인터넷의 활용이 확대되면서 첫째의 포함률 오차는 점차 감소되고 있으나 둘째의 자기선택 오차는 근본적인 문제로 남는다.

인터넷 조사표본의 편향을 가중치로써 제거하는 성향조정 가중법(propensity adjustment weighting)의 이론적 틀은 Rosenbaum & Rubin(1983, 1984)에 의해 정립된 바 있다. 이 방법의 활용에 있어 핵심적 관건은 어떤 사회심리적 요인에 근거하여 가중치를 산출할 것인가에 있으며 최종 관심은 가중치 적용의 결과가 모집단 추정치로서 어느 정도 정확할 것인가에 있다.

이러한 방법을 최초로 선거예측에 적용한 Harris Interactive의 경우 결과는 정확하였지만(Taylor 2000, Taylor et al. 2001), 이들이 어떤 변수들을 성향보정에 사용하였는가는 발표논문에 기술되어 있지 않다. 이러한 유형에 속하는 또 하나의 연구인 Hahs-Vaughn & Onwuegbuzie(2006)도 가중치의 산출과정을 상세히 설명하였으나 어떤 변수들을 가중치 산출에 사용하였는지 명확히 밝히지 않았다.

국내에서도 성향조정 가중법이 적용된 다수의 조사 사례가 보고되었다. 김원용·이흥철(2003)이 2002년 12월의 제 16대 대통령선거에 대한 인터넷 조사자료에 성향조정 가중법을 적용한 예측결과를 보고한 바 있지만 구체적으로 어떤 변수가 성향조정의 기준으로 사용되었는지는 기술되지 않았다. 그러나 이계오·장덕현(2009)은 2007년 12월의 제 17대 대통령 선거에서 인터넷 조사 자료에 성향조정 가중법을 적용한 예측결과를 보고하면서 성향조정 기준변수로 국가보안법 폐지, 대형주택 중과세, 북·미 관계, 대북지원, FTA, 미국 선호도, 일본 선호도, 중국 선호도, 북한 선호도, 이념 성향 등을 활용하였음을 밝혔다. 김영원·배예영(2010)도 2007년 12월의 제 17대 대통령선거에서 인터넷 조사 자료에 성향조정 가중법을 적용한 예측결과를 보고하였는데, 그들은 지역, 성별, 연령, 직업, 교육수준, 가구 소득과 2002년 대선 투표후보, 2004년 국회의원 선거 투표정당, 17대 대선 투표의향, 지지 정당 등의 변수가 성향조정 기준변수로 가중치 산출 모형에 포함되었다고 밝혔다. 이상의 3개 국내 사례에서는 인터넷 조사와 성향조정 가중법에 대한 평가가 모두 긍정적이었다.

성향조정 가중법에 대한 연구자들의 평가가 항상 긍정적인 것은 아니었다. 외국의 경우, 예컨대 Duffy et al.(2005)은 미국 Harris Interactive와 영국의 MORI가 공동으로 영국에서 실시한 면대면 조사와 온라인 조사 자료에 이 방법을 적용하였지만 평가변수가 무엇이냐에 따라 성향조정 가중법의 효과가 다르게 나타났다고 보고하였다. 이들은 인구사회적 변수로서 면대면 조사 자료에는 지역, 사회계층, 차량소유, 나이, 성과 경제활동상태 등에 근거하여, 온라인 조사 자료에는 연령, 성, 지역, 교육수준, 소득수준, 인터넷 사용정도 등에 근거하여 rim(iterative proportional; raking ratio) 가중치를 산출하였다. 그리고 온라인 구매, 개인 정보량에 대한 의견, 위험·사회적 압력·규율 등에 대한 견해 등 5개 문항

을 성향조정 기준변수로 활용하여 가중치를 산출하였다. 평가변수는 투표의사, 정치적 활동성향, 이민에 대한 견해, 콜레스테롤에 대한 태도 등이었고, 성향조정 가중법의 성공여부는 평가변수에 따라 다르게 나타났다고 보고하였다.

Schonlau et al.(2004)에서는 성향조정 가중법이 기대한 효과를 가져 오지 않았다. 이들은 캘리포니아 주민을 대상으로 건강에 관한 RDD 전화조사와 인터넷 조사를 실시하여 결과를 비교하였는데, 성향조정 가중치를 적용했음에도 불구하고 추정치 간의 차이가 37개 문항 중 29개 문항에서 제거되지 않았다. 또 다른 연구인 Schonlau et al.(2009)에서도 성향조정 가중법이 그다지 효과적이지 않은 것으로 나타났다. 이들은 50세 이상을 대상으로 하는 미국인의 건강과 은퇴에 관한 연구(Health and Retirement Study)를 하면서 인터넷 접근 가능자 중에서 표본을 추출하였다. 이들이 알아보고자 한 것은 인터넷 표본이 전체 표본을 대표할 수 있는가, 즉 선택편향(selection bias)이 없어지는가를 보고자 인종, 성별, 연령, 소득, 건강상태에 대한 자기평가, 주택소유여부 등을 토대로 성향조정 가중치를 구하여 적용했지만 추정치 간의 차이가 충분히 감소하지 않았다.

벨기에에서 실시된 연구에서도 성, 연령, 교육수준, 거주지역 및 직업유무 등을 토대로 성향조정 가중치를 구하고 이것을 온라인 응답자 패널 조사결과에 적용했지만, 근로만족도, 정치적 태도, 이민에 대한 태도 등에 관한 온라인 조사와 면접조사 간의 차이가 감소하지 않았다(Loosveldt & Sonck 2008).

이처럼 성향조정 가중법에 대한 평가가 부정적인 경우도 있었으나 몇 개의 부정적 연구 결과들은 주로 인구사회적 속성들에 근거하여 성향조정 가중치를 산출하였다는 점에서 일정한 한계가 있다. 따라서 방법론적으로는 성향조정 가중법을 사용하고 있지만 기존의 인구사회적 속성에 의한 조사자료 가중법 적용과 크게 다를 바 없는 것으로 보인다.

성향조정 가중법을 적용하기 위해서는 확률적 표본추출에 의한 충분한 규모의 준거조사(reference survey)가 요구되는데, 성공적으로 보고된 일부 국내 사례의 경우 이 점에서는 다소 취약한 면이 있는 것이 사실이다. 대부분의 사례에서 준거조사의 크기가 충분하지 않았고 일반적인 한계를 피할 수 없는 할당추출(quota sampling)을 사용하였기 때문이다. 또한 성공적으로 보고된 국내 사례들이 모두 선거예측에 국한되었으므로 일반화에 있어 한계가 있어 보인다.

이런 배경에서 인터넷 조사자료의 보정에 적용될 성향조정 기준변수를 찾고 평가변수를 다양하게 하며, 보다 일반적인 사회조사에 활용 가능한 구체적 성향조정 가중법의 방법론 제시와 평가에 이 연구의 목표를 둔다.

II장에서는 면대면 조사 표본과 인터넷 조사 표본의 차이를 탐색하여 주요 성향적 차이를 밝히고, III장에서는 이를 반영하여 성향조정 가중치를 도출하고 가중치 적용의 결과를 평가한다. IV장에서 성향보정 가중법의 유용성을 논의하고 마지막 V장에서는 이 연구의 활용방안을 제시한다.

II. 면대면 조사 대 인터넷 조사

1. 인터넷 조사 표본 추출

본 연구에서는 인터넷 조사를 위해 자원자들의 참여를 유도하여 패널을 구축하였고 패널에서 일부를 추출하여 조사표본으로 하였다. 원칙적으로 패널의 크기는 작은 것보다 큰 것이 좋겠지만 크기가 가장 중요하다고 보기 어렵다. 조사 응답자로서의 패널 가입자들의 ‘질’, 즉 다각적 측면에서의 대표성도 못지않게 중요하다. 김영원·배예영(2010)의 지적처럼 인터넷에서 자원자를 모집하는 방식보다는 확률표본 응답자에게 패널참여를 권유하여 패널을 만드는 것이 좋을 것이다. 인터넷에서 자원자를 모집하는 경우에도 홍보 창구와 방식에 따라 구성되는 패널에 질적 차이가 있을 것으로 예상된다.

본 연구에서 표본은 모집단과 일치된 주요한 인구·지리적 구성을 갖도록 기획되었다. 즉 광역시·도(16), 동·읍면(2), 성별(2), 연령대(5) 등에서 표본이 모집단과 일치하도록 부집단(subgroup)의 크기를 정하였다(괄호 안 숫자는 범주 수를 나타냄). 따라서 부집단의 총 수는 250개로 다음과 같이 구성되었다.

$$(7\text{개 광역시} + 9\text{개 도}) \times (2\text{개 동} \cdot \text{읍면}) \times (2\text{개 성별}) \times (5\text{개 연령대})$$

이 연구에서 모집된 패널은 2009년 7월 14일에 최종 7,616명으로 주로 통계청 홈페이지를 통해 얻어졌다. 모집기간이 짧았고 홍보 창구도 제한적이었으므로 규모가 크지 않았다. 그나마 인터넷 조사 자원자들이므로 250개의 모든 인구·지리적 부집단에서 모집단 비례 표본을 확보하는 것은 가능하지 않다. 인터넷 사용자 집단은 특히 연령대 분포에서 모집단과 큰 차이가 있기 때문이다. <표 1>은 상주인구(모집단) 대비 자원자 패널의 상대적 비를 본 것이다. 이상적 비는 100%이지만 50대 연령대와 60대 이상 연령대의 상대적 비는 모집단 대비 각각 27%와 6%에 불과하였고, 반면 30대 연령대의 상대적 비는 모집단 대비 175%에 달하였다. 또한 읍면 거주자의 모집단 대비 상대적 비는 58%에 그쳤다.

〈표 1〉 상주인구 대비 자원자 패널의 상대적 비

시·도	서울 1.20, 부산 1.12, 대구 1.20, 인천 0.99, 광주 0.94, 대전 1.68, 울산 1.14, 경기 0.93, 강원 0.97, 충북 0.84, 충남 0.64, 전북 0.89, 전남 0.62, 경북 0.76, 경남 0.79, 제주 0.88
동·읍면	동 1.09, 읍면 0.58
성 별	남자 0.89, 여자 1.11
연령대	20대 이하 1.60, 30대 1.75, 40대 0.73, 50대 0.27, 60대 이상 0.06

목표 표본이 인구·지리적 부집단에서 인구 비례하는 크기를 갖도록 하는 것이 어려우므로, 차선으로 각 인구·지리적 특성 별로 주변적으로 모집단과 일치하는 분포를 갖도록 하였다. 다음은 본 연구에서 고안되고 적용된 표본추출 방법이다.

가중 리스트 계통 추출법(systematic sampling from the weighted list)

1. 패널 리스트를 광역시·도(16), 동·읍면(2), 성별(2), 연령대(5)의 순서로 정렬한다.
2. 정렬된 리스트에 $rim(raking\ ratio)$ 가중법에 의한 가중치를 부여한다. 이때, 가중치의 총계는 패널 리스트의 길이인 N 임을 가정한다. 개체 가중치 w 가 4 또는 5로 설정된 절단값(cut value)보다 크면 절삭 처리한다(예비 단계에서는 절단값을 5로 하였고 최종 단계에서는 절단값을 4로 하였다). 기준집단은 최근 주민등록(상주) 인구이고 일치 기준변수는 광역시·도, 동·읍면, 성별, 연령대 등이다. 적용된 기준 비율은 〈표 2〉와 같다.
3. 0과 가중치 총계의 구간에서 다음 수열 $\{a_m; m = 1, \dots, n\}$ 을 빼낸다.
 - 항 간 간격은 $k = N_0/n$ 이다(등차수열). 여기서 N_0 는 패널 리스트에 부여된 가중치의 총계이고 n 은 목표 표본의 크기이다.
 - 초항 a_1 은 구간 $(0, k)$ 에서 임의로 선정된다(random selection).
 - k 가 절단값보다 작으면 1개 개체가 2회 이상 추출될 가능성이 있으므로 절단값을 하향 조정한다.
4. 각 $m (= 1, \dots, n)$ 에 대하여, 가중치의 부분 합(partial sum)에서 a_m 을 초과하는 인덱스 중에서 가장 작은 것을 찾아낸다.

〈표 2〉 rim 가중법 적용 기준 비율

시·도	서울 0.209, 부산 0.075, 대구 0.052, 인천 0.054, 광주 0.030, 대전 0.031, 울산 0.022 경기 0.221, 강원 0.031, 충북 0.031, 충남 0.039, 전북 0.038, 전남 0.038, 경북 0.054, 경남 0.064, 제주 0.011
동·읍면	동 0.818, 읍면 0.182
성 별	남자 0.491, 여자 0.509
연령대	20대 이하 0.255, 30대 0.221, 40대 0.217, 50대 0.139, 60대 이상 0.168

〈표 3〉 상주인구 대비 추출 표본의 상대적 비

시·도	서울 1.03, 부산 0.97, 대구 1.08, 인천 1.13, 광주 0.90, 대전 1.13, 울산 1.14 경기 0.99, 강원 0.94, 충북 0.97, 충남 0.85, 전북 1.11, 전남 0.95, 경북 0.85, 경남 0.95, 제주 1.00
동·읍면	동 1.03, 읍면 0.85
성 별	남자 0.95, 여자 1.05
연령대	20대 이하 1.16, 30대 1.16, 40대 1.16, 50대 1.07, 60대 이상 0.29

〈표 3〉은 상주인구 대비 추출 표본의 상대적 비를 본 것이다. 〈표 1〉의 상주인구 대비 자원자 패널의 상대적 비에 비해 상당히 개선되었음을 확인할 수 있다. 예컨대 50대 연령대와 60대 이상 연령대가 패널에서는 모집단 대비 상대 비율이 27%와 6%에 불과했으나 표본에서는 각각 107%와 29%로 개선되었고, 읍면 거주자는 모집단 대비 상대 비율이 패널에서는 58%였으나 표본에서는 85%로 개선되었다.

이상의 가중 리스트 계통추출법에 따라 뽑은 표본에서도 주요 인구·지리적 구성에서 모집단과 완전히 일치하지는 않는다. 이는 소수 개체에 지나친 가중치가 붙는 것을 막았기 때문이다. 모집단과 표본의 불일치는 이후 사후 가중치 부여를 통해 해결될 것이다.

앞서 패널 리스트가 광역시·도, 동·읍면, 성별, 연령대의 순서로 정렬되어 있음을 가정한 바 있다. 따라서 목표 응답자의 접촉 실패 또는 비협조 등의 사유가 발생하는 경우 패널 리스트에서 다음 줄에 위치한 예비 응답자로 대체하더라도 최종 표본의 인구·지리적

구성은 목표로부터 크게 벗어나지 않게 된다.

2. 면대면 조사 자료와 인터넷 조사 자료의 비교

지역추출과 조사원 면접 방식에 의한 2009년 사회조사가 7월에 실시되었고 8월 말 분석용 자료로 완결되었다(이하 FTF 표본으로 약칭; 응답자 수 37,049명). 설문지는 문화 및 여가, 복지, 소득 및 소비, 노동, 사회참여 등 5개 영역으로 구성되었다. 조사표를 포함한 사회조사에 대한 세부적 사항은 통계청 사회조사 홈페이지(<http://survey.go.kr/society/>)에서 찾을 수 있다.

조사 이후, 4개 인구·지리적 요인(광역시·도, 동·읍면, 성별, 연령대)에 의한 결합된 가중법(cell weighting)으로 개체별 가중치가 부여되었다. FTF 표본의 cell 가중치는 최소값 0.14, Q1 0.55, 중간값 0.84, Q3 1.25, 최대값 8.97로 요약된 바 범위가 넓고 퍼진 상태였다(가중치는 평균 1이 되도록 척도화되었음). 참고로, FTF 표본에 대한 rim 가중치를 구하여 요약한 결과 최소값 0.29, Q1 0.59, 중간값 0.83, Q3 1.47, 최대값 2.18로 나타났다(가중치는 평균 1이 되도록 척도화되었음). 이처럼 rim 가중치의 산포(범위)가 cell 가중치의 산포(범위)에 비해 작았다. cell 가중치와 rim 가중치의 상관계수는 0.85였다.

자원자 패널과 인터넷을 매체로 한 인터넷 사회조사의 설문지는 문화 및 여가, 소득 및 소비, 사회참여 등 3개 영역으로 구성되었다. 인터넷 조사는 2009년 7월에 실시되었고 곧바로 분석용 자료가 완결되었다(이하 인터넷 표본으로 약칭; 응답자 수 2,903명). 분석자료 개체에 4개 인구·지리적 요인(광역시·도, 동·읍면, 성별, 연령대)에 의한 rim 가중법으로 가중치가 부여되었는데 이때 절단값으로 5를 적용하였다. 인터넷 표본의 rim 가중치는 최소값 0.52, Q1 0.73, 중간값 0.89, Q3 1.05, 최대값 4.58로 나타났다.

이와 같이 FTF 표본에서와 인터넷 표본에서 다른 가중법이 적용된 이유는 표본크기가 현격히 다르기 때문이다. 표본크기가 충분히 크지 않은 경우에는 rim 가중법이 안정적이다.

2009년 통계청 사회조사의 FTF 표본(cell 가중치 적용)과 인터넷 표본(rim 가중치 적용)의 차이를 살펴보기로 한다. 두 표본의 차이를 살펴보는 이유는 인터넷 자원자의 편향성을 파악하기 위해서이다. 주요 사항은 다음과 같다.

1) 교육정도(초등/중학교, 고등학교, 대학교, 대학원 이상)

(단위: 명)

		초등/중학교	고등학교	대학교	대학원 이상	합 계
FTF	빈 도	8905	13379	13352	1413	37049
	%	24.0	36.1	36.0	3.8	100.0
인터넷	빈 도	315	871	1548	169	2903
	%	10.9	30.0	53.3	5.8	100.0
합 계	빈 도	9220	14250	14900	1582	39952
	%	23.1	35.7	37.3	4.0	100.0

FTF 표본에서는 대학/대학원 교육이 39.8%였으나 인터넷 표본에서는 대학/대학원 교육이 59.1%로 19.3%P의 차이를 보였다. 반면 FTF 표본에서는 초등/중학교 교육이 24.0%였으나 인터넷 표본에서는 초등/중학교 교육이 10.9%에 불과하였다. 이와 같이 FTF 표본 응답자에 비해 인터넷 표본 응답자들의 교육수준이 높은 것으로 나타났다(감마계수 0.34).

2) 소비생활만족도(5점 척도: 1=매우 만족, 2=약간 만족, 3=보통, 4=약간 불만족, 5=매우 불만족)

(단위: 명)

		매우만족	약간만족	보통	약간불만족	매우만족	합계
FTF	빈 도	666	4410	17072	11297	3603	37048
	%	1.8	11.9	46.1	30.5	9.7	100.0
인터넷	빈 도	80	520	1135	873	295	2903
	%	2.8	17.9	39.1	30.1	10.2	100.0
합계	빈 도	746	4903	18207	12170	3898	39951
	%	1.9	12.3	45.6	30.5	9.8	100.0

소비생활만족도는 응답자가 자각하는 소비수준인데, 만족 이상의 비율이 FTF 표본에서는 13.7%였으나 인터넷 표본에서는 20.7%로 7.0%P의 차이를 보였다. 이와 같이 FTF 표본 응답자에 비하여 인터넷 표본 응답자들의 자각적 소비수준이 약간 높은 것으로 나타났다(감마계수 0.06).

3) 일반신문을 읽는 빈도(5점 척도: 1=거의 매일, 2=1주일에 3-4회, 3=1주일에 1-2회, 4=2주일에 1회, 5=본 적 없다) (단위: 명)

		거의 매일	1주일에 3~4회	1주일에 1~2회	2주일에 1회	본 적 없다	합 계
FTF	빈 도	8820	3989	4763	2095	17381	37048
	%	23.8	10.8	12.9	5.7	46.9	100.0
인터넷	빈 도	927	352	409	233	982	2903
	%	31.9	12.1	14.1	8.0	33.8	100.0
합 계	빈 도	9747	4341	5172	2328	18363	39951
	%	24.4	10.9	12.9	5.8	46.0	100.0

거의 매일 일반신문을 읽는 응답자의 비율이 FTF 표본에서는 23.8%였으나 인터넷 표본에서는 31.9%로 8.1%P의 차이를 보였다. 반면 일반신문을 본 적 없다는 응답자 비율은 FTF 표본에서는 46.9%였고 인터넷 표본에서는 33.8%로 13.1%P의 차이가 났다. 이와 같이 FTF 표본 응답자에 비하여 인터넷 표본 응답자들이 일반신문을 자주 접하는 것으로 나타났다(감마계수 0.19).

4) 인터넷 신문을 읽는 빈도(5점 척도: 1=거의 매일, ... , 5=본 적 없다) (단위: 명)

		거의 매일	1주일에 3~4회	1주일에 1~2회	2주일에 1회	본 적 없다	합 계
FTF	빈 도	8979	4656	4061	1651	17703	37050
	%	24.2	12.6	11.0	4.5	47.8	100.0
인터넷	빈 도	1215	576	311	125	676	2903
	%	41.9	19.8	10.7	4.3	23.3	100.0
합 계	빈 도	10194	5232	4372	1776	18379	39953
	%	25.5	13.1	10.9	4.4	46.0	100.0

거의 매일 인터넷 신문을 읽는 응답자의 비율이 FTF 표본에서는 24.2%였으나 인터넷 표본에서는 41.9%로 17.7%P의 차이를 보였다. 반면 인터넷 신문을 본 적 없다는 응답자 비율은 FTF 표본에서는 47.8%였고 인터넷 표본에서는 23.3%로 24.5%P의 차이가 났다. 이와 같이 FTF 표본 응답자에 비하여 인터넷 표본 응답자들이 인터넷 신문을 자주 접하는 것으로 나타났다(감마계수 0.39). 그러한 경향이 일반신문의 경우에서도 있었으나 인터넷신문의 경우에 더 강하다고 하겠다.

5) 소득분배에 대한 견해(5점 척도: 1=매우 공평, 2=약간 공평, 3=보통, 4=약간 불공평, 5=매우 불공평) (단위: 명)

		매우 공평	약간 공평	보통	약간불공평	매우불공평	합 계
FTF	빈도	212	1308	9539	16046	9944	37049
	%	0.6	3.5	25.7	43.3	26.8	100.0
인터넷	빈도	26	114	572	1205	986	2903
	%	0.9	3.9	19.7	41.5	34.0	100.0
합 계	빈도	238	1422	10111	17251	10930	39952
	%	0.6	3.6	25.3	43.2	27.4	100.0

소득분배가 매우 불공평하다는 응답자 비율이 FTF 표본에서는 26.8%였으나 인터넷 표본에서는 34.0%로 7.2%P의 차이를 보였다(감마계수 0.13).

6) 계층상승 가능성에 대한 견해(1=매우 높다, 2=비교적 높다, 3=비교적 낮다, 4=매우 낮다, 5=모르겠다) (단위: 명)

		매우 높다	비교적 높다	비교적 낮다	매우 낮다	모르겠다	합 계
FTF	빈도	1133	12794	12665	4228	6228	37048
	%	3.1	34.5	34.2	11.4	16.8	100.0
인터넷	빈도	134	908	1262	405	194	2903
	%	4.6	31.3	43.5	14.0	6.7	100.0
합 계	빈도	1267	13702	13927	4633	6422	39951
	%	3.2	34.3	34.9	11.6	16.1	100.0

계층상승 가능성이 낮다(3:비교적 낮다, 4:매우 낮다)는 응답자 비율이 FTF 표본에서는 45.6%였으나 인터넷 표본에서는 57.5%로 11.9%p의 차이를 보였다. 한편, 모르겠다는 DK(Don't Know) 비율이 FTF 표본에서는 16.8%, 인터넷 표본에서는 6.7%로 인터넷 응답자들이 적극적으로 의사 표현을 하는 것으로 나타났다. 그러나 이 차이는 조사방식의 차이일 수 있다.

이상을 종합 정리하면, FTF 표본 응답자들에 비해 인터넷 표본 응답자들은 교육정도가 높고 소비수준이 높으며 일반신문과 인터넷 신문을 자주 접하는 편이다. 또한 인터넷 표본 응답자들이 상대적으로 더 비판적이고 비판적인 사회의식을 갖는 경향이 있다고 하겠다. 이에 따라 주요 인구·지리적 요인을 통제한다고 하더라도 다수의 사회·문화 영역에서 FTF 조사와 인터넷 조사가 성향적 차이를 보일 개연성이 충분히 있는 것으로 보인다.

III. 성향조정 가중법

1. 성향조정 로짓모형 가중법

Rosenbaum & Rubin(1983, 1984)이 제안한 성향조정 가중법(propensity adjustment weighting)은 관측연구(observational study)에서 집단 간에 존재하는 성향적 차이를 제거하여 집단 간 공정한 비교를 가능하게 한다. 이 방법을 요약하면 다음과 같다.

집단 A와 집단 B에서 표본을 취하였고 관심변수 Y 와 함께 다수의 배경변수(공변량) X_1, \dots, X_p 가 측정되었다고 하자. 집단 A의 Y 와 집단 B의 Y 를 공정하게 비교하기 위하여 로짓모형

$$\log \frac{P(A | x_1, \dots, x_p)}{P(B | x_1, \dots, x_p)} = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p \quad (1)$$

를 활용해 개체별로 A에 속할 확률 $P(A | x_1, \dots, x_p)$ 와 B에 속할 확률 $P(B | x_1, \dots, x_p)$ 를 추정한다(두 확률의 합은 1이다). 이런 확률들의 추정값을 성향점수(propensity score)라고 한다.

성향점수의 활용 방법은 크게 다음 세 가지이다(D'agostino 1998). 첫째 방법은 짝짓기 기법(matching technique)으로 성향점수가 비슷한 A 개체 1개와 B 개체 1개 또는 m 개(= 2, 3, ...)를 쌍으로 묶어 쌍 내에서 Y 를 비교하고, 다수의 쌍에서 나오는 비교결과들

을 종합한다. 둘째 방법은 층화(stratification) 또는 가중법(weighting method)으로, 1) 집단 A와 집단 B의 합 집단(pooled group)의 차원에서 A 개체는 $P(A | x_1, \dots, x_p)$ 의 역수로 가중하고 B 개체는 $P(B | x_1, \dots, x_p)$ 의 역수로 가중한다. 또는 2) 집단 A를 기준으로 하는 경우에는 A 개체는 가중하지 않고(가중치=1) B 개체는

$$\frac{P(A | x_1, \dots, x_p)}{P(B | x_1, \dots, x_p)} \tag{2}$$

로 가중한다. 안정적인 가중치 부여를 위해 성향점수를 구간화하기도 하나 꼭 그렇게 해야 하는 것은 아니다(구간화를 하면 가중치 분포의 범위를 줄임으로써 안정적인 결과를 만들어내겠지만 반대급부도 예상된다). 마지막 방법은 회귀 수정(regression adjustment)이다.

인터넷 조사 결과의 보정이 목적인 경우 앞의 두 번째 방법을 적용한다. 그때 A는 준거 조사(reference survey)인 FTF 표본이고 B는 인터넷 표본이다. 즉, 지지자 T를 FTF 표본이면 1, 인터넷 표본이면 0으로 두고 $P(T=1 | x_1, \dots, x_p)$ 에 대한 로짓모형 추정값 p를 구한다. 이어서 FTF 표본 개체에는 성향조정 가중치 1을 부여하고 인터넷 표본 개체에는 성향조정 가중치 p/q 를 부여한다($q = 1 - p$).

각 표본 개체들에 인구·지리적 구성의 목적으로 이미 cell 또는 rim 가중치가 부여된 경우 앞의 로짓모형 적합 시 그런 가중치가 반영되어야 하며, 기존 cell/rim 가중치에 성향조정 가중치를 곱하는 가중치 업데이트 작업이 실행된다.

로짓모형의 적용에서 핵심 이슈는 성향보정의 기준변수(공변량, 배경변수)의 선정에 있다. 성향조정 기준변수의 선정 시 다음 몇 가지 사항을 고려하여야 할 것으로 생각된다.

- 1) 성향조정 기준변수는 변화하지 않는 내재적 근본 성향을 반영하여야 하며, 변하기 쉬운 표면적 변수는 피해야 한다. 예컨대 정치여론 조사에서 ‘한반도 통일의 당위성’에 대한 의견은 내재적이거나 ‘6자 회담 참석’에 대한 의견은 표면적이다.
- 2) 쉽게 질문할 수 있고 쉽게 응답할 수 있어야 한다. 성향조정 기준변수는 보정의 기반이 되므로 측정오차가 개입될 소지가 적어야 한다. 어려운 질문이나 부담이 되는 응답에는 심각한 측정오차가 개입될 위험이 있다.
- 3) 그 변수에서 집단 간 유의한 차이가 있어야 하고 그 차이가 이해 가능해야 한다. FTF 표본 응답자와 인터넷 표본 응답자는 인터넷 접근과 활용 빈도에서 차이가 있을 것이므로 그런 차이를 반영할 수 있어야 한다.

II장에서 FTF 표본과 인터넷 표본의 차이를 탐색한 바를 참고하고 성향조정 기준변수의 선정에 관한 위 관점들을 지침으로 하여, 본 연구에서는 다음 변수들을 성향조정 기준변수를 선정하고 범주화하였다.

- 1) 광역시·도(16개 층) = 1(서울), 2(부산), ... , 16(제주)
- 2) 동·읍면(2개 층) = 1(동), 2(읍면)
- 3) 남·여(2개 층) = 1(남), 2(여)
- 4) 연령(5개 층) = 1(20대-), 2(30대), 3(40대), 4(50대), 5(60대+)
- 5) 교육정도 (4개 수준) = 1(초/중), 2(고), 3(대학), 4(대학원)
- 6) 소비생활만족도 (2개 수준) = 1(만족), 2(보통/불만족)
- 7) 소득분해 견해 (5개 수준, 리커트 척도)
- 8) 일반신문 보기 (2개 수준) = 1(본다), 2(보지 않는다)
- 9) 인터넷 신문 보기 (5개 수준, 리커트 척도)

첫 4개 변수는 인구·지리적 구성의 대표성을 위해 쓰인 바 있고 이미 cell 및 rim 가중치로써 균형화되도록 조치된 바 있다. 그러나 rim 가중화에서 절단값을 적용한 바 있고 다른 변수들과 결합하여 새로운 효과를 만들어 낼 가능성이 있으므로 모형에 포함하였다.

본 연구의 경우, FTF 표본은 37,049명의 응답자로 구성되었고 인터넷 표본은 2,903명의 응답자로 구성되었으므로 크기에 있어 큰 차이가 있다. 성향조정 로짓모형 (1)의 적합에 앞선 사전 균형화(prior balancing) 조치로서 인터넷 표본을 12.76배(=37,049/2,903)로 부스팅(boosting)하여 크기를 FTF 표본과 맞추었다. 부스팅의 결과로 모형 (1)의 상수항 β_0 는 달라지나 설명항의 계수들 β_1, \dots, β_p 는 변하지 않는다. 그러나 부스팅으로 로짓모형의 보다 안정적인 추정이 가능할 것으로 기대한다.

이상 9개 변수를 설명변수로 성향조정 로짓모형이 적합되었다. 로짓모형에서 회귀계수는 다음과 같다(회귀계수에 대한 표준오차 및 t 값, p-값 등은 부스팅의 영향으로 축소, 증대, 축소되었을 것이므로 보고하지 않았다. 부스팅 없이도 분석표본의 총 크기가 크므로 대부분 계수들이 유의하게 나타나게 된다).

이 모형에서의 계수들은 FTF로 이끌어질 성향을 나타내므로, 양이면 FTF일 가능성으로 기울고 음이면 인터넷일 가능성으로 기운다. 계수의 절대값 크기는 그런 가능성의 강도이다.

시도코드1	0	연령대3	-0.29
시도코드2	-0.18	연령대4	-0.64
시도코드3	-0.18	연령대5	-1.10
시도코드4	-0.21	교육정도1	0
시도코드5	-0.21	교육정도2	-0.26
시도코드6	-0.14	교육정도3	-0.58
시도코드7	-0.20	교육정도4	-0.51
시도코드8	-0.07	일반신문1	0
시도코드9	-0.14	일반신문2	0.33
시도코드10	-0.17	인터넷신문1	0
시도코드11	-0.14	인터넷신문2	0.09
시도코드12	-0.19	인터넷신문3	0.51
시도코드13	-0.20	인터넷신문4	0.49
시도코드14	-0.13	인터넷신문5	1.06
시도코드15	-0.17	소득분배1	0
시도코드16	-0.19	소득분배2	0.31
동읍면1	0	소득분배3	0.44
동읍면2	-0.11	소득분배4	0.25
성별1	0	소득분배5	-0.07
성별2	-0.22	소비생활만족도1	0
연령대1	0	소비생활만족도2	0.39
연령대2	-0.00		

여기서는 4개 인구·지리적 변수 외 성향조정 기준변수들의 효과를 살펴보기로 한다.

- 1) 가장 큰 효과를 보인 변수는 인터넷 신문 보기이다(1: 거의 매일, 2: 1주일에 3-4회, 3: 1주일에 1-2회, 4: 2주일에 1회, 5: 보지 않는다). 인터넷 신문을 보지 않는 응답자는 거의 매일 보는 응답자에 비해 FTF 대 인터넷의 오즈(odds)가 2.89배 (=exp(1.06))나 된다. 오즈 비가 1보다 큰 그 정도로 증대됨을 의미한다.

- 2) 교육정도가 대학 및 대학원인 응답자는 초등/중학교인 응답자에 비해 FTF 대 인터넷의 오즈가 0.56배(=exp(-0.58)) 내지 0.60배(=exp(-0.51))가 된다. 오즈 비가 1 보다 작음은 그 정도로 축소됨을 의미한다.
- 3) 소득분배의 공평성에 대한 견해가 보통인 응답자는 매우 공평하다고 생각하는 응답자에 비하여 FTF 대 인터넷의 오즈가 1.55배(=exp(0.44))가 된다(그 정도로 증대된다). 매우 불공평하다고 생각하는 응답자에 비한 오즈도 이와 비슷하다.
- 4) 소비생활만족도가 보통/불만족/매우 불만족인 응답자는 만족/매우 만족인 응답자에 비해 FTF 대 인터넷의 오즈가 1.48배(=exp(0.39))가 된다(그 정도로 증대된다).
- 5) 일반신문을 보지 않는 응답자는 보는 응답자에 비해 FTF 대 인터넷의 오즈가 1.39배(=exp(0.33))가 된다(그 정도로 증대된다).

이와 같은 기준변수들의 효과는 가법적으로 작용하므로, 인터넷 신문을 보지 않고 교육정도가 대학이며 소득분배의 공평성에 대해 보통이라고 생각하며 소비생활만족도가 보통/불만족/매우 불만족이며 일반신문을 보지 않은 응답자는, 인터넷 신문을 거의 매일 보고 교육정도가 초등/중학교이며 소득분배의 공평성에 대해 매우 공평하다고 생각하며 소비생활만족도가 만족/매우 만족이며 일반신문을 보는 응답자에 비해 FTF 대 인터넷의 오즈가

$$2.89 \times 0.56 \times 1.55 \times 1.48 \times 1.39 \text{배, 즉 } 5.16 \text{배}$$

가 된다. 그 정도로 증대되는 것이다.

성향조정 가중치 부여로 인하여 다시 인구·지리적 구성이 흐트러지므로 이를 재조정할 필요가 있는데(Lee & Valliant 2009), 본 연구에서는 이를 위하여 성향조정 가중치가 붙은 인터넷 표본에 rim 가중법을 적용하였으며 절단값은 5로 설정하였다. 이것은 일종의 보정(calibration) 절차이다(Deville & Sarndal 1992). 그러나 본 연구에서는 성향조정 로짓모형의 적합에 앞서 FTF 표본과 인터넷 표본을 모두 인구·지리적 구성에 맞게 가중화하였고 성향점수 모형에 4개의 기본 인구·지리적 변수를 포함시킨 바 있으므로, 보정효과는 크지 않았다. 최종 rim 가중법에서 절단값을 설정함으로써 일부 개체 가중치가 절단된 것 외에는 큰 변화가 없었음을 알 수 있다. <표 4>가 본 연구에 적용된 인터넷 표본에 대한 가중치 산출 과정을 보여준다.

〈표 4〉 인터넷 표본에 대한 가중치 산출 과정

	원자료	→	단계 1	→	단계 2	→	단계 3
가중하는 이유	해당 없음		인구·지리적 구성		성향조정		인구·지리적 구성
가중치 산출 방법	해당 없음		rim weighting		propensity score model		rim weighting
부여가중치	1	→	w1	→	$w2 := w1 \times \frac{p}{q}$	→	w3

2. 성향조정 rim 가중법

성향조정 로짓모형 가중법으로 보정을 하는 대신에 같은 기준변수들로 인터넷 표본의 rim 가중화를 통해 집단 간에 존재하는 인구·지리적, 성향적 차이를 제거함으로써 집단 간의 공정한 비교를 위한 토대를 마련할 수 있다.

1절에서와 같이, 다음 변수들을 성향조정 기준변수로 선정하였다.

- 1) 광역시·도(16개 층) = 1(서울), 2(부산), ... , 16(제주)
- 2) 동·읍면(2개 층) = 1(동), 2(읍면)
- 3) 남·여(2개 층) = 1(남), 2(여)
- 4) 연령(5개 층) = 1(20대-), 2(30대), 3(40대), 4(50대), 5(60대+)
- 5) 교육정도 (4개 수준) = 1(초/중), 2(고), 3(대학), 4(대학원)
- 6) 소비생활만족도 (2개 수준) = 1(만족), 2(보통/불만족)
- 7) 소득분배 견해 (5개 수준, 리커트 척도)
- 8) 일반신문 보기 (2개 수준) = 1(본다), 2(보지 않는다)
- 9) 인터넷 신문 보기 (5개 수준, 리커트 척도)

rim 가중법을 적용하려면 성향조정 기준변수 9개 각각에 대하여 모집단 비율이 있어야 한다. 그 가운데 4개의 인구·지리적 변수에 대해서는 모집단 비율이 있지만 그 밖의 5개 변수에 대해서는 모집단 비율이 없다. 차선책으로, 4개의 인구·지리적 변수에 의해 cell 가중화된 FTF 표본에서 그 밖의 5개 변수에 대한 모집단 비율을 추정하였다.

본 연구에 적용된 rim 가중법에 의한 인터넷 표본의 성향조정 과정을 단계별로 기술하면 다음과 같다.

단계 1: FTF 표본을 4개 인구·지리적 cell 가중치를 부여하여 그 밖의 5개 성향조정 기준변수의 분포를 구한다. cell 가중화 대신 rim 가중화를 할 수도 있다.

단계 2: 9개 변수의 기준 분포에 맞춰 인터넷 표본을 rim 가중화한다. 지나치게 큰 가중치에 대하여는 절단값을 적용할 수 있다.

성향조정 로짓모형 가중법과 마찬가지로 이 방법도 일종의 보정 절차로 볼 수 있으므로 (Deville & Samdal 1992) 두 방법의 적용 결과 간에 큰 차이가 없을 것으로 예상할 수 있다. 전자에 비해 단순한 것이 장점이다.

3. 주요 평가문항의 성향보정 결과

2009년 사회조사의 전 문항(인터넷 조사 설문지)에 걸친 본격적인 성향보정의 결과는 다음 장에서 평가할 것이다. 이 절에서는 성향조정 로짓모형 가중법에 의한 성향보정과 rim 가중법에 의한 성향보정의 결과를 몇 개의 주요 평가문항에서 살펴보기로 한다.

이 절에서 평가문항으로 선정한 설문 및 척도는 다음과 같다.

- TV 평일 시청(시간)
- TV 토요일 시청(시간)
- TV 일요일 시청(시간)
- 해외여행 경험(지난 1년간 비율, %)
- 전반적 만족감(5점 리커트 척도 평균, 1 = 매우 만족, ... , 5 = 매우 불만족)
- 여가활용 만족도(5점 리커트 척도 평균, 1 = 매우 만족, ... , 5 = 매우 불만족)
- 장래소득 기대(5점 리커트 척도 평균, 1 = 크게 증가, ... , 5 = 크게 감소)
- 주관적 건강평가(5점 리커트 척도 평균, 1 = 매우 좋다, ... , 5 = 매우 나쁘다)
- 스트레스 인식(4점 리커트 척도 평균, 1 = 매우 많다, ... , 4 = 거의 없다)
- 생활시간 태도(5점 리커트 척도 평균, 1 = 매우 바쁘다, ... , 5 = 전혀 그렇지 않다)

<표 5>는 성향조정 로짓모형(PL) 가중법에 의한 성향보정의 결과이다. 이 표에서 각 열은 FTF 표본과 인터넷 표본으로부터 다음 방법으로 얻은 수치이다. 이 중에서 [5]의 방법에 대하여는 2절에서 자세히 기술한 바 있다.

[1] FTF 표본 원자료(FTF raw)

[2] FTF 표본 인구·지리적 cell 가중치 적용(FTF cell)

[3] 인터넷 표본 원자료(인터넷 raw)

- [4] 인터넷 표본 인구·지리적 rim 가중치 적용(인터넷 rim)
- [5] 인터넷 표본 성향조정 로짓모형 가중치 적용(인터넷 PL)

〈표 5〉를 보면 TV 평일 시청시간에 대한 FTF raw는 2.891시간, FTF cell은 2.841시간으로 별 차이가 없다. 그런데 인터넷 raw는 2.528시간이고 인터넷 rim은 2.674시간으로 FTF raw나 FTF cell에 비해 작다. 인터넷 응답자의 TV 시청 시간이 적은 이유는 인터넷 사용과 관련 있을 것으로 추측된다. 그러나 인터넷 표본에 성향점수모형 가중치를 적용한 결과인 인터넷 PL이 2.760시간으로 나타나 충분하지는 않지만 FTF cell과의 차이가 상당 부분 좁혀진 것으로 나타났다. 성향조정 가중법의 긍정적 역할을 볼 수 있는 경우이다.

사회조사의 경우 FTF 표본이나 인터넷 표본의 어떤 표본에서든 인구·지리적 가중치를 적용하여 활용하므로 이제부터는 더 이상 [1]의 FTF raw와 [3]의 인터넷 raw를 언급하지 않기로 한다. 즉, [2]의 FTF cell과 [4]의 인터넷 rim 간 차이를 살펴보고 [5]의 인터넷 성향조정 추정치가 [4]의 인터넷 rim보다 [2]의 FTF cell에 근접하는가를 볼 것이다.

〈표 5〉 성향조정 로짓모형(PL) 가중법에 의한 성향보정의 결과

	[1]* FTF raw	[2] FTF cell	[3] 인터넷 raw	[4] 인터넷 rim	[5] 인터넷 PL	비고**
TV 평일	2.891	2.841	2.528	2.674	2.760	[5]
TV 토	3.608	3.600	3.518	3.611	3.663	[4]
TV 일	3.925	3.922	3.848	3.917	3.998	[4]
해외여행 경험	12.2%	13.6%	15.5%	15.3%	13.3%	[5]
전반적 만족감	3.186	3.170	3.182	3.201	3.265	[4]
여가활용 만족도	3.134	3.128	2.970	2.988	3.065	[5]
장래소득 기대	2.935	2.875	2.490	2.591	2.626	[5]
주관적 건강평가	2.831	2.789	2.602	2.697	2.761	[5]
스트레스 인식	2.541	2.518	2.306	2.337	2.335	-
생활시간 태도	2.653	2.611	2.493	2.585	2.617	[5]

* [1] FTF(면대면) 표본 원자료(FTF raw)
 [2] FTF(면대면) 표본 인구·지리적 cell 가중치 적용(FTF cell)
 [3] 인터넷 표본 원자료(인터넷 raw)
 [4] 인터넷 표본 인구·지리적 rim 가중치 적용(인터넷 rim)
 [5] 인터넷 표본 성향조정 로짓모형 가중치 적용(인터넷 PL)

** [4]와 [5] 중에서 [2]에 더 가까운 추정치.

TV 토요일 시청시간과 일요일 시청시간에서 [2]의 FTF cell과 [4]의 인터넷 rim은 거의 차이가 없다. 그런데 성향보정을 하게 되면서 시청시간 추정치가 늘어나 오차가 발생하게 되는 것을 볼 수 있다. 성향조정 로짓모형에 의한 가중법이 부정적인 경우이다.

해외여행 경험 비율에서 [2]의 FTF cell은 13.6%인 반면 [4]의 인터넷 rim은 15.3%인데 이는 인터넷 응답자들이 교육정도와 소비수준이 높은 데 기인하는 것으로 보인다. 인터넷 표본추정치를 성향보정하면 13.3%로 [2]의 FTF cell의 13.6%에 근접하므로 이 경우는 성향조정 로짓모형 가중법에 의한 보정이 매우 효과적임을 보여준다.

전반적 만족감(경제, 직업, 건강 등)에서 [2]의 FTF cell은 3.17점이고 [4]의 인터넷 rim은 3.201점으로 인터넷 응답자들이 다소 더 불만족하는 편이다. 그런데 인터넷 표본추정치를 성향보정하면 3.265점으로 [2]의 FTF cell에서 더욱 멀어지는 결과가 초래된다. 이 경우는 성향조정 로짓모형 가중법에 의한 보정이 제 역할을 하지 못 함을 보여준다.

그밖에 여가활용 만족도, 장래소득 기대, 주관적 건강평가, 생활시간 태도 등을 살펴보면 성향조정 로짓모형 가중법에 의한 보정이 어느 정도의 역할을 하는 것으로 볼 수 있다. 다만, 스트레스 인식에 있어서는 [2]의 FTF cell은 2.518점이고 [4]의 인터넷 rim은 2.337점으로 인터넷 응답자들이 더 스트레스를 느끼는 편인데 인터넷 표본을 성향보정하여 얻는 추정치는 2.335점으로 별로 달라지지 않는 것으로 나타났다. 종합해 보면, 10개 문항으로 성향조정 로짓모형 가중법에 의한 보정 방법을 평가한 결과 6개 문항에서는 긍정적으로, 3개 문항에서는 부정적으로, 1개 문항에서는 부정도 긍정도 아닌 것으로 나타났다.

향후 여러 사례에 2009년 사회조사 자료를 활용하여 인터넷 표본을 성향보정하기 위해서는 가급적 성향조정 기준변수의 수를 작게 하는 것이 좋을 것이다. 이에 따라 성향조정 모형에서 덜 중요한 것으로 나타나는 소비생활 만족도(2개 수준)와 일반신문 보기(2개 수준)를 제외한 7개 변수를 설명변수로 하는 로짓모형을 만들고 이를 기초로 인터넷 표본을 가중화하는 성향보정을 하여 보았다. <표 6>이 그 결과이다.

전반적으로 볼 때, 기준변수의 수가 9개인 <표 5>와 비교하여 주목할 만한 차이는 나타나지 않았다. 그러나 세부적인 수치 변동이 일부 있어서, 10개 문항의 문항별 평가를 종합하면 7개 문항에서는 긍정적으로, 2개 문항에서는 부정적으로, 1개 문항에서는 부정도 긍정도 아닌 것으로 나타났다.

<표 7>과 <표 8>은 rim 가중법에 의한 성향보정의 결과로 기준변수가 각각 9개와 7개인 경우를 보여준다.

〈표 6〉 성향조정 로짓모형(PL) 가중법에 의한 성향보정의 결과 기준변수를 7개로 줄인 경우

	[1]* FTF raw	[2] FTF cell	[3] 인터넷 raw	[4] 인터넷 rim	[5] 인터넷 PL	비고**
TV 평일	2.891	2.841	2.528	2.674	2.756	[5]
TV 토	3.608	3.600	3.518	3.611	3.668	[4]
TV 일	3.925	3.922	3.848	3.917	3.988	[4]
해외여행경험	12.2%	13.6%	15.5%	15.3%	13.7%	[5]
전반적만족감	3.186	3.170	3.182	3.201	3.205	-
여가활용만족도	3.134	3.128	2.970	2.988	3.020	[5]
장래소득기대	2.935	2.875	2.490	2.591	2.623	[5]
주관적건강평가	2.831	2.789	2.602	2.697	2.734	[5]
스트레스 인식	2.541	2.518	2.306	2.337	2.353	[5]
생활시간태도	2.653	2.611	2.493	2.585	2.624	[5]

대체로 〈표 7〉은 〈표 5〉와 유사하고 〈표 8〉은 〈표 6〉과 유사한 것으로 나타났다. 따라서 성향조정 방법에 의해 결과가 영향을 받기 보다는 기준변수의 선정에 의한 영향이 더 큰 것으로 보인다.

지금까지의 연구 내용을 요약하면서 이 장을 마무리하면 다음과 같다.

- 1) 통계청 사회조사 인터넷 표본의 성향조정 기준변수를 9개로 한 경우의 결과와 7개로 한 경우의 결과는 별 차이가 없다. 그러나 보다 안정적인 결과도출을 위해서 전 문항에 걸친 성향조정에서는 보정기준 변수를 9개로 둘 것이다. 보정변수를 9개로 설정한 이유는 다음 절에서 기술하기로 한다.
- 2) 성향조정 로짓모형 가중법에 의한 성향보정의 결과와 rim 가중법에 의한 성향보정의 결과는 별 차이가 없었다.
- 3) 인터넷 표본의 성향조정 가중법에서 성공의 관건은 기준변수의 선정에 있다.
- 4) 통계청 사회조사 인터넷 표본의 성향조정 결과, 자원자 패널을 활용한 인터넷 조사가 면접원 조사를 일부 보완할 수 있는 대안적 방법으로 보인다.

〈표 7〉 성향조정 rim 가중법에 의한 성향보정의 결과

	[1]* FTF raw	[2] FTF cell	[3] 인터넷 raw	[4] 인터넷 rim	[5] 인터넷 P.rim	비고**
TV 평일	2.891	2.841	2.528	2.674	2.763	[5]
TV 토	3.608	3.600	3.518	3.611	3.664	[4]
TV 일	3.925	3.922	3.848	3.917	3.990	[4]
해외여행 경험	12.2%	13.6%	15.5%	15.3%	13.1%	[5]
전반적 만족감	3.186	3.170	3.182	3.201	3.273	[4]
여가활용 만족도	3.134	3.128	2.970	2.988	3.073	[5]
장래소득 기대	2.935	2.875	2.490	2.591	2.631	[5]
주관적 건강평가	2.831	2.789	2.602	2.697	2.764	[5]
스트레스 인식	2.541	2.518	2.306	2.337	2.332	-
생활시간 태도	2.653	2.611	2.493	2.585	2.616	[5]

〈표 8〉 성향조정 rim 가중법에 의한 성향보정의 결과: 기준변수를 7개로 줄인 경우

	[1]* FTF raw	[2] FTF cell	[3] 인터넷 raw	[4] 인터넷 rim	[5] 인터넷 P.rim	비고**
TV 평일	2.891	2.841	2.528	2.674	2.757	[5]
TV 토	3.608	3.600	3.518	3.611	3.669	[4]
TV 일	3.925	3.922	3.848	3.917	3.990	[4]
해외여행 경험	12.2%	13.6%	15.5%	15.3%	13.6%	[5]
전반적 만족감	3.186	3.170	3.182	3.201	3.207	-
여가활용 만족도	3.134	3.128	2.970	2.988	3.025	[5]
장래소득 기대	2.935	2.875	2.490	2.591	2.627	[5]
주관적 건강평가	2.831	2.789	2.602	2.697	2.735	[5]
스트레스 인식	2.541	2.518	2.306	2.337	2.352	[5]
생활시간 태도	2.653	2.611	2.493	2.585	2.623	[5]

IV. 성향조정 가중법의 유용성 검토

앞 장에서는 인터넷 자원자 패널 표본의 편향을 제거하기 위해 성향조정 가중치를 생성 시키고 예비적으로 10개의 대표적 평가변수에 적용한 바 있다. 이 장에서는 2009년 통계청 사회조사에 포함된 총 106개 문항(성향조정 기준변수 제외)에 가중치를 적용해 그 결과를 분석한다. 검토될 가중법은 9개 변수 성향조정 로짓모형 가중법(이하 LPW9으로 약칭)과 7개 변수 성향조정 로짓모형 가중법(이하 LPW7으로 약칭)이다. 성향조정 rim 가중법은 로짓모형 가중법과 거의 같은 결과를 내기 때문에 제외시켰다.

성향조정 가중법의 적용 결과에 대한 평가는 준거조사인 면대면 조사를 기준으로 했을 때 성향조정 가중치 적용 추정치가 기본가중치 적용 추정치에 비해 기준치에 가까운 정도로 하였다. 그런데 문항별로 응답항의 수가 다르기 때문에 편차의 크기를 집계하는 데 어려움이 있다. 따라서 각 문항에서 가장 대표적인 응답 항을 정하고 그것을 경계로 범주화하여 답 항 수를 2로 통일하였다. 예컨대 독서의 경우는 1년 동안 읽은 서적의 수가 3권 이상인 경우와 그 미만인 경우에 대한 퍼센트를 산출하였다.

결과를 요약하면, LPW9 적용 추정치가 기본가중치 적용 추정치에 비해 준거조사 추정치에 더 가깝게 나타난 문항은 106개 문항 중 84개였고, LPW7의 경우는 106개 문항 중 83개였다. 퍼센트로 환산하면 각각 79%와 78%이다. 따라서 대략 5개 문항 중 4개 문항에서 성향조정 가중법이 효용이 있는 것으로 볼 수 있다.

더 나아가, 향상된 정확도가 어느 정도인가를 분석하였다. 기준으로서 기본가중치를 적용하였을 때의 인터넷 조사 추정치와 준거조사 추정치 간 차이를 산출한 결과 평균절대값 편차는 5.9%P였다. 그런데 LPW9을 적용한 경우 평균절대값 편차는 4.6%P로 감소하였다. 따라서 정확도 향상은 1.3%P라고 할 수 있다. 이는 상대적으로 22%(=1.3/5.9)에 해당하는 수치이다. LPW7의 경우는 평균절대값 편차가 4.8%P였다. 따라서 정확도 향상은 1.1%P이고 이는 상대적으로는 19%(=1.1/5.9)에 해당하는 수치이다.

다음으로 성향조정 가중법의 효용이 조사 분야에 따라 어떻게 나타나는지를 알아보았다. <표 9>는 2009년 통계청 온라인 사회조사에 포함된 조사 분야이다. 매체접촉형태 11개 문항에서는 평균적으로 기본가중치 적용 추정치와 준거조사 추정치 차이가 10.9%P인데 LPW9과 준거조사 추정치 차이가 8.0%P로 나타나 상대적으로 20% 정도의 정확도 향상이 있었다. 이 분야에서 LPW7은 LPW9보다 다소 못한 것으로 나타났다. 특이한 분야는 소득·소비·근로 분야의 12개 문항인데 평균적으로 기본가중치 적용 추정치와 준거조사 추정치 차이가 6.8%P인데 LPW9이나 LPW7이 이 차이를 전혀 좁히지 못하였음을 볼 수 있다.

〈표 9〉 2009년 통계청 사회조사 분야

분야	문항 수	문항 내용
매체 접촉형태	11	TV 시청시간, 독서시간, 잡지류, 교양서적, 직업관련 서적, 생활관련 서적, 기타
관광, 여가, 문화활동	31	골프장, 국내관광여행 정도, 해외여행경험, 관광, 어학연수, 해외여행유무, 레저시설 이용횟수, 관광명소, 온천장 및 스파, 스키장, 해수욕장, 삼림욕장, 놀이공원, 수용장 워터파크, 기타, 문화예술 및 스포츠, 음악, 연극, 무용, 영화, 박물관, 미술관, 스포츠, 주말 여가활동, 여가활동
인간관계 및 생활의례	8	생활의례, 인간관계 및 예절, 민속놀이, 전통예술, 전통복장, 전통건축, 전통무예,
소득, 소비, 근로	12	소득만족도, 장래소득기대, 계층의식, 계층이동, 근로시간
사회관계	7	아플 경우의 친구, 돈 빌릴 경우, 낙심할 경우
봉사, 자선활동	18	후원금, 자원봉사활동 등
생활 만족도	19	건강, 노동, 교육 등

〈표 10〉 인터넷 조사와 준거 조사 간 평균 절대값 편차

(단위: %P)

조사분야	문항 수	기본 차이	LPW9	LPW7
매체 접촉형태	11	10.9	8.0	8.3
관광, 여가, 문화활동	31	5.0	2.9	3.2
인간관계 및 생활의례	8	6.7	4.7	5.3
소득, 소비, 근로	12	6.8	7.6	7.3
사회관계	7	9.0	7.8	8.2
봉사, 자선활동	18	2.1	1.9	1.8
생활 만족도	19	6.1	4.8	5.0

V. 요약 및 결론

인터넷 조사의 경우 전통적인 조사방법의 한계를 극복할 수 있는 대안으로 주목받아 왔지만 지금까지 표본의 대표성 문제 때문에 활용이 제한되어 왔다. 그럼에도 불구하고 인터넷 보급의 확대 그리고 전통적인 조사방법의 문제점 때문에 인터넷 조사방법에 대한 관심

도는 더욱더 높아지고 있다. 본 연구는 이러한 맥락에서 인터넷 조사의 문제점을 극복하는 방안을 찾아보고 검증해 보았다.

2009년 통계청 먼대면 사회조사에 준거하여 온라인 사회조사에 성향조정 가중법을 적용한 결과 다음 사실을 발견하였다.

- 1) 성향조정 가중법이 5개 평가문항 중 4개 문항 정도에서 추정치의 정확도 향상에 기여하였다. 정확도 향상분은 상대적으로 20% 정도로 나타났다.
- 2) 성향조정 가중법의 편차감소 효과는 소득·소비·근로 분야에서는 잘 나타나지 않았다. 이와 같이 성향조정 가중법의 효과는 분야에 따라 다르다.
- 3) 성향조정 로짓모형 가중법과 성향조정 rim 가중법 간에는 차이가 없었다.
- 4) 광역시·도, 동·읍면, 성별, 연령대, 교육정도, 소비생활만족도, 소득분배 견해, 일반신문 보기, 인터넷신문 보기 등 9개 변수를 쓰는 것이 소비생활만족도, 일반신문 보기 등 2개 변수를 줄여 7개 변수를 쓰는 것보다 나은 것으로 나타났지만 그 차이는 미미하였다.

이와 같은 연구결과를 토대로 자원자 패널에 의한 인터넷 조사방법의 활용가능성에 대해 결론을 내리면 다음과 같다.

첫째, 본 연구의 결과는 성향조정 가중법의 유용성을 보여준다. 본 연구에서는 비록 그 크기가 충분하지는 않지만 80% 정도의 문항에서 긍정적으로 기여하는 것으로 나타났기 때문이다. 지금까지의 인터넷 성향조정 가중법 사례들이 일부 제한된 변수들에 기초하였다는 점을 감안하면 본 연구가 과거의 연구들보다 성향가중 모형의 유용성과 한계를 더 잘 보여주었다고 생각한다. 본 연구에서는 다양한 분야의 106개 문항을 평가변수로 사용했고 조정 변수로는 기본 인구적 속성 외에 매체접촉 행태, 소득분배 견해 등 사회심리적 변수들을 두루 사용했기 때문이다. 따라서 이번 연구에서 편차 감소폭이 크지 않았지만 이것은 앞으로 좀더 강력한 기준변수를 선택함으로써 개선될 수 있을 것으로 본다.

둘째, 성향조정 가중법에 포함되는 성향조정 기준변수는 평가변수에 따라 즉, 적용분야에 따라 달리 선택되어야 한다. 본 연구에서 성향조정 가중법의 효과는 평가변수 분야에 따라 다르게 나타났다. 따라서 이 점은 적용 분야에 따라 기준변수를 달리 선택해야 한다는 것을 의미한다. 예컨대 정치 분야, 과학 분야, 소비행동 분야 등과 같이 특정 분야별로 성향조정 기준변수들을 발굴할 필요가 있음을 의미한다.

셋째, 성향조정 가중법에서 rim 가중법과 로짓모형 가중법 간에 차이가 없다는 점은 성

향조정 가중법의 적용성을 더욱 높여준다. 로짓모형 가중법을 하려면 무엇보다도 준거조사의 원자료가 필요하기 때문이다. 그러나 rim 가중치를 적용하기 위해서는 원자료가 필요없이 요약 통계치로도 충분하다. 이 점은 센서스 또는 기타 대규모 조사결과에서 얻은 많은 자료를 보다 쉽게 활용할 수 있다는 것을 의미한다. 아울러 성향조정 로짓모형 가중법의 경우에는 성향조정 기준변수가 하나의 준거조사에 모두 포함되어 있어야만 활용이 가능했다. 이 점 역시 기준변수의 선택을 제한하는 큰 요인이 되고 있다. 그러나 rim 가중치를 사용할 경우 여러 개의 조사에서 나온 결과를 활용할 수 있게 된다. 즉 기준변수의 선정이 보다 확대되었다는 것을 의미하며, 이것이 성향조정 가중법의 활성화에 기여할 것으로 보인다.

본 연구의 활용 방안을 직접적인 측면과 부차적인 측면에서 생각해 보면 다음과 같다.

- 1) 직접적 활용: 매년 거듭되는 통계청 사회조사에서 현재의 면접조사 방식을 인터넷 자원자 조사로 점차 대체해 가는 방안이다. 인터넷 조사가 접촉이 어려운 사회구성원에 접근할 수 있고 사생활 침해를 염려하는 응답자들에 심리적 부담을 줄여 줄 것이다. 그러나 인터넷 자원자 조사로 면접 조사를 완전히 대체할 수는 없다. 성향조정의 준거가 없어지기 때문이다.
- 2) 간접적 활용: 성향조정의 준거가 되는 2009년 사회조사 FTF 표본 원자료 및 기준변수의 분포를 공개하여 통계청 및 기타 국가나 민간기관이 시행하는 각종 인터넷 조사 표본의 성향조정에 활용하는 방안이다. 이런 방식의 활용을 위해서는 본 연구에서 성향조정의 기준으로 제시한 7개 또는 9개 문항이 인터넷 조사 설문지에 포함되어야 한다.
 - ① 성향조정 로짓모형 가중법을 활용하려면 2009년 사회조사 FTF 표본 7개 또는 9개 변수 원자료가 필요하고 이를 인터넷 조사 표본에 붙여 성향조정 로짓모형을 만들어 성향조정 가중치를 생성하면 된다.
 - ② 성향조정 rim 가중법을 활용하기 위해서는 2009년 사회조사 cell 가중치 부여 FTF 표본으로부터 7개 또는 9개 변수의 주변 분포 비율을 제공받아 rim 가중법을 적용하면 된다.

또다른 활용 방안은 혼합방식(mixed mode) 조사에서 조사방식에 따른 차이를 규명하고 방식별 추정값들을 통합하여 하나의 추정치를 제공하는 데 있어 본 조사의 방법론을 응용하는 것이다. 향후 사회여건의 변화에 따라 혼합방식의 조사가 불가피해지는 만큼 본 연구

에서 수행한 방법론의 부가적 활용도가 있을 것으로 기대한다.

향후에도 꾸준히 성향조정 기준변수에 대한 연구가 필요하다. 성향조정 가중법의 유용성을 좀더 높이기 위해서는 본 연구에서 선정한 9개 기준변수 외에 어떤 다른 기준변수들이 가능한지 찾아보는 노력이 필요하다. 이 점은 기준변수가 조사 분야에 따라 다를 수 있다는 점에서 더욱 중요해진다. 이 사례에서 미흡하였던 소득·소비·근로 등의 영역에 작동하는 기준변수를 찾아보는 노력이 필요하다.

참고문헌

- 김영원·배예영. 2010. “인터넷 선거여론조사 가중치보정을 위한 성향점수의 활용.” 《한국통계학회 논문집》 17(1): 55-66.
- 김원용·이홍철. 2003. “웹조사의 모집단대표성 확보를 위한 성향가중 모형의 적합성 검증.” 《방송연구》 2003년 여름호: 143-166.
- 이계오·장덕현. 2009. “인터넷 선거조사에서 성향가중모형 적용사례.” 《조사연구》 10(3): 21-36.
- D'agostino, R. B. Jr. 1998. “Tutorial in Biostatistics: Propensity Score methods for Bias Reduction in the Comparison of a Treatment to a Ton-randomized Control Group.” *Statistics in Medicine* 17: 2265-2281.
- Deville, J. C., and C. E. Samdal, 1992. “Calibration Estimators in Survey Sampling.” *Journal of the American Statistical Association* 87: 376-382.
- Duffy, B., K. Smith, G. Terhanian, and J. Bremer. 2005. “Comparing Data from Online and Face-to-face Surveys.” *International Journal of Market Research* 47(6): 615-639.
- Hahs-Vaughn, D. L. and A. J. Onwuegbuzie. 2006. “Estimating and Using Propensity Score Analysis with Complex Samples.” *The Journal of Experimental Education* 75(1): 31-65.
- Lee, S., and R. Valliant. 2009. “Estimation for Volunteer Panel Web Surveys Using Propensity Score Adjustment and Calibration Adjustment.” *Sociological Methods and Research* 37(3): 319-343.
- Loosveldt G. and N. Sonck. 2008. “An Evaluation of the Weighting Procedures for an Online Access Panel Survey.” *Survey Research Methods* 2(2): 93-105.
- Rosenbaum, P., and D. B. Rubin. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” *Biometrika* 70: 41-55.
- Rosenbaum, P., and D. B. Rubin. 1984. “Reducing Bias in Observational Studies Using

- Subclassification on the Propensity Score.” *Journal of the American Statistical Association* 79: 516–524.
- Schonlau, M., A. Soest, A. Kapteyn, and M. Couper. 2009. “Selection Bias in Web Surveys and the Use of Propensity Scores.” *Sociological Methods & Research* 37(3): 291–318.
- Schonlau, M., K. Zapert, L. P. Simon, K. H. Sanstad, S. M. Marcus, J. Adams, H. Kan, R. Turner, and S. Berry. 2004. “A Comparison between Responses from a Propensity-weighted Web Survey and an Identical RDD Survey.” *Social Science Computer Review* 22(1): 128–138.
- Taylor, H. 2000. “Does Internet Research Work? Comparing Online Survey Result with Telephone Survey.” *International Journal of Market Research* 42: 58–63.
- Taylor, H., J. Bremer, C. Overmeyer, J.W. Siegel, and G. Terhanian. 2001. “The Record of Internet Based Opinion Polls in Predicting the Results of 72 Races in the November 2000 US Elections.” *International Journal of Market Research* 43: 12.

[접수 2010/3/31, 수정 2010/5/10, 게재확정 2010/5/20]