

연구논문

패널 이탈과 조건화의 영향:

KBS-MBMR의 2007년 대통령 선거 패널조사 사례

Impact of Panel Attrition and Conditioning:

The Case of KBS-MBMR's 2007 Korean Presidential Election Panel Survey

허명회* · 김지연** · 이양훈***

Myung - Hoe Huh · JeeYeon Kim · RyangHoon Lee

패널조사가 시점 간 변화를 알아내는 데 있어 매우 유용한 조사방법론이기는 하지만 조사차수가 넘어갈수록 패널표본이 모집단 대표성을 잃어 간다는 단점이 있다. 이러한 표본의 왜곡은 패널 이탈과 조건화에 기인하는 것으로 알려져 있다. 본 연구는 4개월 동안 5회 측정된 KBS-MBMR의 2007년 대통령 선거 패널조사에서의 이탈과 조건화를 통계적 관점에서 검토해 본 것이다. 주요 결과는 다음과 같다. 첫째, 패널 이탈 성향은 여자가 남자보다, 20대 이하와 60대 이상이 30대, 40대, 50대보다, 선호(적합/지지) 후보를 묻는 질문에 대해 “없다/모른다/무응답”으로 대답한 사람(=DK 응답자)들이 선호후보 밝힘 응답자들보다 컸다. 둘째, 패널조사라는 반복측정이 선호후보 밝힘을 응답하도록 밀어붙이는 조건화 효과가 있음을 확인하였다. 셋째, 이러한 패널 이탈과 조건화에도 불구하고, DK 응답을 제외하는 경우 반복조사가 지지후보 응답에 영향을 주지 않았고 최종 투표의향에도 영향을 주지 않았다.

주제어: 패널 이탈, 패널 조건화, 선거조사

Even though panel surveys are very useful in estimating the change between time points, they suffer from sample distortion as survey rounds proceed due to panel attrition and conditioning. This study is to report the statistical aspects of KBS-MBMR's five-rounds panel survey for the 2007 Presidential Election of Korea. Main findings are as follows. First, the attrition propensities were higher among women, among the young and the old aged, and among "None"/"Don't Know"/"Won't Say"(DK) respondents for the question asking preferred candidate. Second, there existed the conditioning effect that pushes the respondents to choose one candidate in hurry. Third, repeated measurements of the survey did not influence

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

E-mail: stat420@korea.ac.kr

** 밀워드브라운미디어리서치 사회여론조사본부 본부장/상무

*** 밀워드브라운미디어리서치 사회여론조사본부 1팀장/차장

panel respondents on the choice of preferred candidate, if DK respondents are set aside, and on the final voting intention of panel respondents.

Key words: panel attrition, panel conditioning, election survey

I. 연구 배경

종단적 연구(longitudinal research)에서 패널조사가 차지하고 있는 위치는 독보적이다. 특히 시간의 흐름에 따라 여론의 향방과 후보자에 대한 지지 정도가ダイナ믹하게 변하는 선거조사 영역에서 더욱 유용하다. 일반적인 추세조사(tracking survey)를 통해서 특정 시점 간에 여론변화가 있을 경우 그 원인을 규명하기 어렵지만, 패널조사에서는 개인 수준의 태도변화를 포착할 수 있고 그 이유를 물을 수 있기 때문이다. 이러한 장점에도 불구하고 일반적으로 패널조사는 이탈(attrition)과 조건화(conditioning)로 대표성이 저해될 수 있다는 문제점을 갖는 것으로 생각된다(Bartels 1999; Clinton 2001). 우리는 본 논문에서 우리나라의 선거 패널조사에서 그런 일반적 문제점이 실제 존재하는지를 검토하고자 한다.

패널조사에 관한 기존 연구들은 패널구성의 편향성과 응답자 조건화를 우려하면서도 주목적 정보의 도출에 있어 주목할 만한 문제가 야기되지 않을 수 있음을 비치고 있다. 예컨대, Dennis & Li(2003)는 유지 패널과 이탈 패널을 비교한 결과 두 패널이 서로 비슷한 태도, 가치, 행동패턴을 보였음을 보고하였다. 컴퓨터 및 인터넷 사용에 있어서 두 패널 간 일정한 차이가 발견되었지만 그 차이의 상당부분이 인구적 요인의 차이로 설명되었던 것이다. 특히 정치적 태도에서는 유지 패널과 이탈 패널이 유사하였다. Taylor et al.(1996)은 영국의 선거 패널조사에서 1차 조사 참여 이후 2차 조사의 거절은 정치에 대한 지적 수준이나 흥미 정도와 관련이 있었음을 보고하였고, Loosveldt & Carton(1997)도 벨기에 총선 패널 조사에서 유사한 결론을 얻었다. Das et al.(2007)은 네덜란드 가구 패널조사에서 지식과 관련된 일부 문항에서만 패널 조건화의 효과가 나타났을 뿐 다른 문항들에서는 패널 조건화나 이탈로 인한 영향이 나타나지 않았음을 보고하였다. Clinton(2001)은, 매주 실시되는 Knowledge Networks 사의 온라인 패널조사에서 지속적으로 참여하는 패널과 전체 패널의 비교를 통하여 패널 이탈 및 조건화의 영향이 거의 없었음을 보고하였다. Zargosky & Rhoton(1999)도 National Longitudinal Survey에서 패널 이탈로 인한 '심각한 문제'가 발생하지 않았음을 보고하였다. Sturgis et al.(2007)은 패널조사에 있어서 특정 주

제에 관련된 질문이 패널 응답자에게 해당 주제에 대한 관심을 유발함으로써 이후의 조사에서 지식수준과 태도에서 다소의 차이가 생기게 되었지만 패널의 이탈로 인한 영향은 거의 없었음을 보고하였다.

선거 패널조사 관련 국내 연구로서, 김장수(2007)와 이내영·이현우·김장수(2007)은 동아시아 연구원, 한국리서치, 중앙일보, SBS가 컨소시엄으로 2006년 지방선거를 위해 서울, 부산, 광주, 충남 4개 지역을 대상으로 실시한 총 4회 패널조사의 과정과 성과를 보고하면서, 선거 패널조사의 의의와 설계, 패널조사로 살펴본 유권자 표심의 변화와 역동성, 그리고 이에 대한 정치학적 해석 등을 다루었다. 또한 이현우·권혁용(2008)은 동아시아 연구원, 한국리서치, 중앙일보, SBS 컨소시엄이 2007년 대통령선거를 위해 전국규모로 실시한 총 6회 패널조사의 성과를 보고하면서, 2007년 선거의 사회경제적 특성과 정치행태, 미디어 효과와 정치참여, 투표 결정요인 등을 분석한 바 있다.

II. KBS-MBMR의 2007년 대통령 선거 패널조사

한국방송 KBS는 2007년 12월의 대통령 선거를 위한 패널조사 용역을 밀워드브라운미디어리서치(MBMR)와 체결하여 다음 일정으로 총 5차에 걸친 조사를 하였다.¹⁾

- 1차 조사: 2007년 8월 15일-17일
- 2차 조사: 2007년 9월 27일-29일
- 3차 조사: 2007년 10월 25일-27일
- 4차 조사: 2007년 11월 30일-12월 2일
- 5차 조사: 2007년 12월 11일-12일

이를 위한 패널구축 조사는 RDD(Random Digit Dialing, 임의번호걸기) 시간균형할당 추출로 함으로써 전화번호부 할당추출에 비해 대표성이 높은 패널 풀을 확보하고자 하였다(허명희 외 2006, 강현철 외 2008). 패널구축 진행은 8월 10일(금)부터 8월 14일(화)까지 5일 간 진행되었으며, 낮 시간과 저녁 시간, 평일과 휴일로 구분하여 7회까지 재통화(call-back)를 실시하였다. <표 1>에서 볼 수 있듯이 최초통화에서는 성공률(응답률)이 6.9%였지만 7차례의 재통화를 실시한 결과 성공률을 16.0%까지 올릴 수 있었다.

1) 3차와 4차 사이인 11월 7일에 긴급현황에 대한 간이조사가 있었고 선거일 당일 패널들의 실제 투표를 알아보기 위한 간이조사가 있었으나 이 연구에 포함하지 않기로 한다.

〈표 1〉 재통화 차수별 패널구축 성공률(응답률) 변화 추이

(단위: %)

구 분	첫 통화	1차 재통화	2차 재통화	3차 재통화	4차 재통화	5차 재통화	6차 재통화	7차 재통화
부재중/통화중	64.2	46.6	37.7	32.0	26.9	22.3	19.1	17.2
조사거절	28.9	42.9	50.1	54.6	58.6	62.5	65.3	66.8
패널구축 성공	6.9	10.5	12.1	13.4	14.5	15.2	15.7	16.0
합 계	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

통상적으로 패널조사는 중도 이탈로 인하여 조사 차수가 지남에 따라 표본 수가 줄어들고 패널 조건화가 심해지기 마련이다. 그러나 조사목적이 선거인만큼 차수가 넘어갈수록 중요도가 커지므로 조사의 크기가 줄어들지 않도록 할 필요가 있어 2차 조사부터 5차 조사까지 각 시기(round)에서 패널표본을 보충하도록 조사를 설계하였다. 다만 보충 패널추출은 1차 조사의 주패널 확보단계에서 부재중/관심없음 등의 이유로 선발되지 않은 응답자 풀에서 하기로 하였다. 패널 이탈이 접촉의 어려움 또는 흥미 결여로 생기는 결과로 생각되기 때문이다. 이렇게 함으로써 실질적으로 응답률을 높이는 결과가 되기도 한다. 패널 이탈률을 감소시키기 위해 KBS 뉴스보도를 통한 참여의식 고취, 패널구축과 실제조사의 분리, 응답부담 최소화, 적절한 사례금 지급 등 제반 조치를 취하였다.

〈표 2〉는 다섯 번의 라운드에서 확보된 패널과 함께 조사된 응답자 수를 보여준다. 1차 라운드에서 2,000명을 확보하고 조사하였으나, 2차 조사에서는 이 중 1,734명(86.7%)으로부터 응답을 받았고 266명(13.3%)으로부터는 응답받지 못하였다. 그러나 2차 라운드에서 새로 428명을 확보함으로써 패널 손실을 충분히 보충하였다. 3차 라운드에서는 208명을, 4차 라운드에서는 210명을, 5차 라운드에서는 219명을 추가로 확보하여 누적 패널은 총 3,065명에 이르렀다. 최종 5차 조사는 총 누적 패널을 대상으로 실시되었고 이 중 2,580명이 응답하였다(응답률 84.2%).

만약 패널 확보가 1차 라운드에 그쳤고 각 조사에서 응답한 패널원만 다음 조사로 넘겼다면 최종 5차 조사의 응답자 수는 1,380명이 되었을 것이다. 이는 최초 패널 2,000명의 69.0%에 해당한다(〈표 3〉을 참조).

〈표 4〉는 차수별 기존 패널의 응답률을 보여 준다. 응답률이 2차 조사 86.7%, 3차 조사 82.4%, 4차 조사 81.7%, 5차 조사 83.0% 등으로 3차 조사 이후 82~83% 수준에서 안정되었음을 볼 수 있다. 이는 4차 라운드와 5차 라운드에서 패널 보충이 없었다라든가 최종 5차 조사에서 2,000명 이상의 응답자를 확보할 수 있었을 것임을 의미한다.

〈표 2〉 패널 확보 시기와 조사 차수별 응답자 표본의 크기

(단위: 명)

	확보	1차 조사	2차 조사	3차 조사	4차 조사	5차 조사
1차 라운드	2,000	2,000	1,734	1,631	1,632	1,650
2차 라운드	428		428	369	339	361
3차 라운드	208			208	183	179
4차 라운드	210				210	171
5차 라운드	219					219
총 합	3,065	2,000	2,162	2,208	2,364	2,580

〈표 3〉 1차 라운드 확보 패널의 후속조사 완결 응답자 수

(단위: 명)

	확보	1차 조사	2차 조사	3차 조사	4차 조사	5차 조사
1차 라운드	2,000	2,000	1,734	1,527	1,435	1,380
	%	100	86.7	76.4	71.8	69.0

〈표 4〉 2차 조사부터 5차 조사까지의 재접촉률

(단위: 명)

	2차 조사	3차 조사	4차 조사	5차 조사
기존 패널의 크기	2,000	2,428	2,636	2,846
응답자 수	1,734	2,000	2,154	2,361
%	86.7	82.4	81.7	83.0

Ⅲ. 패널 이탈과 조건화: 1차 라운드 확보 패널

이 절에서는 1차 라운드 확보 패널에 국한하여 다섯 차례에 걸친 조사에서 총 참여횟수를 인구적 요인에 따라 비교하고 패널 이탈과 조건화의 기본적 양태를 살펴보고자 한다.

〈표 5〉는 1차 라운드에서 확보된 패널의 후속조사(2차, 3차, 4차, 5차) 참여횟수를 권역별로 살펴 본 것이다. 권역별로 참여횟수 분포가 다소 차이는 있으나 특별히 유의하지는 않았다(카이제곱 27.7 (자유도 24), p -값 0.272).

〈표 5〉 1차 라운드 확보 패널원의 권역별 후속조사 참여횟수 분포

	표본 수	후속조사 참여횟수					합계
		0회	1회	2회	3회	4회	
서울	422	7%	7%	4%	12%	70%	100%
인천/경기	532	4%	6%	6%	12%	72%	100%
대전/충남/충북	199	6%	9%	5%	12%	68%	100%
광주/전남/전북	228	9%	4%	5%	15%	66%	100%
대구/경북	196	4%	6%	7%	12%	70%	100%
부산/울산/경남	332	7%	7%	7%	15%	65%	100%
대구/경남	91	9%	7%	7%	10%	68%	100%
전체	2,000	6%	7%	6%	13%	69%	100%

- 카이제곱 27.7(자유도 24), p-값 0.272

〈표 6〉 1차 라운드 확보 패널원의 남녀별 후속조사 참여횟수 분포

	표본 수	후속조사 참여횟수					합계
		0회	1회	2회	3회	4회	
남자	974	4%	5%	6%	11%	73%	100%
여자	1,026	8%	7%	6%	14%	65%	100%
전체	2,000	6%	7%	6%	13%	69%	100%

- 카이제곱 19.9(자유도 4), p-값 0.001

〈표 6〉은 성별로 후속조사 참여횟수 분포를 살펴본 것으로 남녀 간 차이가 유의한 것으로 나타났다(카이제곱 19.9(자유도 4), p-값 0.001). 남자는 모두 참여한 비율이 73%였으나 여자는 65%에 그쳐 8%p의 차이가 나는 등 남자의 참여도가 여자의 참여도에 비하여 큰 것으로 나타났다. 일반적으로 여자가 남자에 비해 재택률이 크다는 점을 고려한다면 접촉 용이도 외에 패널조사 참여에 영향을 주는 요인이 있었음을 알 수 있다.

〈표 7〉은 1차 라운드 확보 패널의 후속조사 참여횟수 분포를 연령대별로 살펴본 결과인데 연령대 간 차이가 통계적으로 유의하게 나타났다(카이제곱 73.8(자유도 16), p-값 0.000). 모두 참여한 비율이 60대 이상에서 58% 그리고, 20대 이하에서는 60%로, 30대의 76%와 40대의 76%에 비해 작았다. 50대에서는 모두 참여한 비율이 72%였다.

〈표 7〉 1차 라운드 확보 패널원의 연령대별 후속조사 참여횟수 분포

	표본 수	후속조사 참여횟수					합계
		0회	1회	2회	3회	4회	
20대-	401	8%	9%	6%	17%	60%	100%
30대	480	3%	5%	5%	11%	76%	100%
40대	467	5%	4%	4%	11%	76%	100%
50대	302	5%	4%	7%	12%	72%	100%
60대+	350	9%	11%	8%	14%	58%	100%
전체	2,000	6%	7%	6%	13%	69%	100%

- 카이제곱 73.8(자유도 16), p-값 0.000

〈표 8〉 1차 라운드 확보 60대 이상 패널원의 남녀별 후속조사 참여횟수 분포

60대+	표본 수	후속조사 참여횟수					합계
		0회	1회	2회	3회	4회	
남자	154	3%	6%	7%	15%	70%	100%
여자	196	14%	16%	9%	13%	49%	100%
전체	350	9%	11%	8%	14%	58%	100%

- 카이제곱 25.3(자유도 4), p-값 0.000

〈표 8〉은 참여도가 가장 떨어지는 60대(이상) 각 그룹을 성별로 나누어 후속조사 참여횟수 분포를 본 교차표이다. 남자는 모두 참여한 비율이 70%였으나 여자는 49%에 그쳐 21%p의 상당히 큰 차이를 보였다. 그러나 60대 이상 남자의 조사 참여도는 다른 연령대 남녀와 큰 차이가 나지 않았다.

이제부터는 패널조사의 동태적 측면을 살펴보기로 한다. 〈표 9〉는 1차 조사와 2차 조사에서 선호후보 밝힘²⁾ 응답과 없다/모른다/무응답(DK)³⁾ 응답 그리고, 조사불가 간 전환행렬(transition matrix)을 보여주는데, 1차 조사 선호후보 밝힘 응답자 중 2차 조사에서 선호

2) 1차 조사부터 3차 조사까지는 적합후보를 물어 보았으며 4차 조사와 5차 조사에서는 지지후보를 물었으나 이 절에서는 '적합'과 '지지'를 구분하지 않고 '선호'로 통용하기로 한다.

3) 여기서 '지지후보가 없다, 모르겠다, 말하지 않겠다(무응답)' 등을 묶어 편의상 DK 응답으로 지칭할 것이다. 따라서 이하의 'DK 응답'은 응답자가 어떤 태도도 갖지 않음을 의미하는 것은 아니다.

〈표 9〉 1차 조사에서 2차 조사로의 전환: 선호후보 밝힘 - 없다/모른다(DK) - 조사불가 (단위: 명)

		2차 조사			합계
		선호후보 밝힘	DK 응답	조사불가	
1차 조사	선호후보 밝힘	1,288(79%)	137(8%)	197(12%)	1,622(100%)
	DK 응답	208(55%)	101(27%)	69(18%)	378(100%)
	조사불가	-	-	-	-
	전체	1,496(75%)	238(12%)	266(13%)	2,000(100%)

〈표 10〉 2차 조사에서 3차 조사로의 전환: 선호후보 밝힘 - 없다/모른다(DK) - 조사불가 (단위: 명)

		3차 조사			합계
		선호후보 밝힘	DK 응답	조사불가	
2차 조사	선호후보 밝힘	1,275(85%)	68(5%)	153(10%)	1,496(100%)
	DK 응답	127(53%)	57(24%)	54(23%)	238(100%)
	조사불가	91(34%)	13(5%)	162(61%)	266(100%)
	전체	1,493(75%)	138(12%)	369(13%)	2,000(100%)

후보 밝힘 응답자로 남는 비율 79%에 비교하여, 1차 조사 DK 응답자 중 2차 조사에서 DK 응답자로 남는 비율은 27%로 상당히 적었다. 선거일이 가까워질수록 선거구도가 명확해지고 캠페인 활동이 가열되므로 유권자 관심이 증가할 것이다. 그러나 DK 응답 비율이 짧은 시일동안 이렇게 급격히 줄어든 것은, 조심스럽지만, 패널 조건화가 작용한 결과일 것으로 해석하고자 한다.⁴⁾ 또한, 1차 조사 선호후보 밝힘 응답자 중 2차 조사에서 조사불가자로 남는 비율이 12%였으나 1차 조사 DK 응답자 중 2차 조사에서 조사불가자로 남는 비율은 18%로 상대적으로 컸다. 이것은 패널 이탈의 일부 원인이 패널원의 DK 태도에 있음을 의미한다.

〈표 10〉은 2차 조사와 3차 조사에서 선호후보 밝힘 응답과 DK 응답 그리고, 조사불가 간 전환행렬을 보여준다. 2차 조사 선호후보 밝힘 응답자 중 3차 조사에서 선호후보 밝힘

4) 이것이 조건화의 결과가 아닌 자연스런 과정인지를 명확히 밝힐 방법은 없다. 그러나 이것이 조건화의 결과가 아니라면 1차 라운드 확보 패널의 최종 조사 DK 비율이 같은 시점에서 처음으로 확보된 5차 라운드 확보 패널의 DK 비율과 유사할 것을 기대해야 할 것이다. 그러나 전자가 후자보다 훨씬 작았다. 5절 참조.

〈표 11〉 3차 조사에서 4차 조사로의 전환: 선호후보 밝힘 - 없다/모른다(DK) - 조사불가 (단위: 명)

		4차 조사			
		선호후보 밝힘	DK 응답	조사불가	합계
3차 조사	선호후보 밝힘	1,336(90%)	70(5%)	87(6%)	1,493(100%)
	DK 응답	84(61%)	31(23%)	23(17%)	138(100%)
	조사불가	94(26%)	17(5%)	258(70%)	369(100%)
전체		1,514(76%)	118(6%)	368(18%)	2,000(100%)

〈표 12〉 4차 조사에서 5차 조사로의 전환: 선호후보 밝힘 - 없다/모른다(DK) - 조사불가 (단위: 명)

		5차 조사			
		선호후보 밝힘	DK 응답	조사불가	합계
4차 조사	선호후보 밝힘	1,383(91%)	44(3%)	87(6%)	1,514(100%)
	DK 응답	69(59%)	38(32%)	11(9%)	118(100%)
	조사불가	98(27%)	18(5%)	252(69%)	368(100%)
전체		1,550(78%)	100(5%)	350(18%)	2,000(100%)

응답자로 남는 비율이 85%로 더 커졌고, 2차 조사 DK 응답자 중 3차 조사에서 DK 응답자로 남는 비율은 24%로 더 작아졌다. 또한 2차 조사 선호후보 밝힘 응답자 중 3차 조사에서 조사불가자로 남는 비율이 10%였으나 2차 조사 DK 응답자 중 3차 조사에서 조사불가자로 남는 비율은 23%로 그 차이가 더 커졌다. 조사불가자의 61%는 조사불가자로 남았다.

〈표 11〉에서 3차 조사에서 4차 조사로의 전환 양태를 보면, “선호후보 밝힘 → 선호후보 밝힘”의 비율과 “조사불가 → 조사불가”의 비율이 더욱 커졌음을 볼 수 있다. “DK 응답 → DK 응답” 비율은 전과 같았지만, “선호후보 밝힘 → 선호후보 밝힘”의 비율에 비하면 상대적으로 매우 작아졌다. 패널 이탈과 조건화가 가속화된 것이다.

〈표 12〉에서 4차 조사에서 5차 조사로 전환되는 양태를 살펴보면, 3차 조사에서 4차 조사로 전환되는 양태와 거의 비슷하다.

이를 종합해 보면 조사 참여에 영향을 주는 요인은 성과 연령, 그리고 지난 시기의 응답 동일 것으로 유추할 수 있다. 다음 절에서 로지스틱 회귀모형을 통하여 이를 확인해 보기로 한다.

IV. 무엇이 후속조사 참여에 영향을 주는가?

〈표 13a〉는 1차 라운드 확보 패널의 2차 조사 참여에 대한 로지스틱 회귀모형의 요약이다. 연령이 매우 유의한 요인으로 꼽혔으며(카이제곱 31.9 (자유도 4), p-값 0.000, 성과 1차 조사 선호후보 밝힘 여부도 유의한 요인으로 꼽혔다. 〈표 13b〉는 유의한 요인들의 로지스틱 회귀모형 계수인데, 각 요인 범주의 계수 간 차이가 연령대에서 가장 크게 나타났다. 즉 60대의 계수가 -0.361, 30대의 계수가 0.570이었는데, 이는 60대의 조사 참여 오즈(odds)가 30대의 조사 참여 오즈의 0.3944배[$=\exp(-0.361-0.570)$]였음을 뜻한다. 1차 조사 DK 응답자의 2차 조사 참여 오즈는 선호후보 밝힘 응답자의 2차 조사 참여 오즈의 0.707배였고 여자는 남자의 0.771배였다. 그만큼 DK 응답자와 여자의 2차 조사 참여 가능성이 작은 것이다.

〈표 14a〉는 1차 라운드 확보 패널의 3차 조사 참여에 대한 로지스틱 회귀모형의 요약이다. 2차 조사 결과(선호후보 밝힘/DK/조사불가)가 가장 유의한 요인으로 꼽혔으며[카이제곱 271.9 (자유도 2), p-값 0.000], 성과 연령도 유의한 요인으로 꼽혔다. 〈표 14b〉는 유의한 요인들의 로지스틱 회귀모형 계수인데 2차 조사에서 조사불가자의 3차 참여 오즈가 선호후보 밝힘 응답자의 3차 참여 오즈의 0.076배에 불과한 것으로 나타났다. 이는 선행조사에서의 조사불가자를 후속 조사에 참여시키기가 매우 어려움을 의미한다. 여전히 여자의 참여율이 남자에 비해 낮았으며, 20대와 60대의 참여율이 낮았고, 30대와 40대가 높게 나타났다.

〈표 15a〉는 1차 라운드 확보 패널의 4차 조사 참여에 대한 로지스틱 회귀모형의 요약이다. 3차 조사 결과가 가장 유의한 요인으로 꼽혔고 연령도 유의하였으나 성은 더 이상 유의하지 않았다. 〈표 15b〉는 유의한 요인들의 로지스틱 회귀모형 계수를 본 것으로 3차 조사에서 조사불가자의 4차 참여 오즈는 더욱 낮아져 선호후보 밝힘 응답자 오즈의 0.027배로 떨어졌다. 20대와 60대의 참여율이 낮았고 30대, 40대, 50대의 참여율이 높게 나타났다. 이런 양상은 5차 조사 참여에서도 마찬가지였다(〈표 16a〉와 〈표 16b〉 참조).

〈표 13a〉 2차 조사 참여에 대한 로지스틱 회귀모형 요약

요인	자유도	카이제곱	p-값
권역	6	11.6	0.071
성	1	3.7	0.054
연령	4	31.9	0.000
1차 조사 선호후보 밝힘	1	4.8	0.028

<표 13b> 2차 조사 참여에 대한 로지스틱 회귀모형 계수(일부)

유의요인	계수	표준오차	exp(계수)	p-값
성: 남자	0.000		1.000	
여자	-0.261	0.135	0.771	0.054
연령: 20대-	0.000		1.000	
30대	0.570	0.209	1.769	0.006
40대	0.483	0.206	1.621	0.019
50대	0.527	0.237	1.694	0.026
60대	-0.361	0.189	0.698	0.057
1차 조사: 선호후보 밝힘	0.000		1.000	
DK 응답	-0.347	0.158	0.707	0.028

<표 14a> 3차 조사 참여에 대한 로지스틱 회귀모형 요약

요인	자유도	카이제곱	p-값
권역	6	6.6	0.359
성	1	10.6	0.001
연령	4	13.0	0.011
2차 조사 결과	2	271.9	0.000

<표 14b> 3차 조사 참여에 대한 로지스틱 회귀모형 계수(일부)

	계수	표준오차	exp(계수)	p-값
성: 남자	0.000		1.000	
여자	-0.428	0.132	0.652	0.001
연령: 20대-	0.000		1.000	
30대	0.639	0.202	1.894	0.002
40대	0.459	0.197	1.582	0.020
50대	0.185	0.213	1.204	0.384
60대+	0.132	0.197	1.141	0.503
2차 조사: 선호후보 밝힘	0.000		1.000	
DK 응답	-0.932	0.181	0.394	0.000
조사불가	-2.580	0.157	0.076	0.000

〈표 15a〉 4차 조사 참여에 대한 로지스틱 회귀모형 요약

요인	자유도	카이제곱	p-값
권역	6	4.2	0.645
성	1	2.3	0.126
연령	4	16.3	0.003
3차 조사 결과	2	498.6	0.000

〈표 15b〉 4차 조사 참여에 대한 로지스틱 회귀모형 계수(일부)

	계수	표준오차	exp(계수)	p-값
연령: 20대-	0.000		1.000	
30대	0.640	0.231	1.897	0.006
40대	0.664	0.229	1.942	0.004
50대	0.669	0.255	1.953	0.009
60대	0.092	0.227	1.097	0.683
3차 조사: 선호후보 밝힘	0.000		1.000	
DK 응답	-1.161	0.260	0.313	0.000
조사불가	-3.595	0.162	0.027	0.000

〈표 16a〉 5차 조사 참여에 대한 로지스틱 회귀모형 요약

요인	자유도	카이제곱	p-값
권역	6	2.1	0.911
성	1	0.0	0.848
연령	4	11.9	0.018
4차 조사 결과	2	498.6	0.000

〈표 16b〉 5차 조사 참여에 대한 로지스틱 회귀모형 계수(일부)

	계수	표준오차	exp(계수)	p-값
연령: 20대-	0.000		1.000	
30대	0.451	0.228	1.570	0.048
40대	0.697	0.233	2.007	0.003
50대	0.702	0.263	2.018	0.007
60대	0.318	0.230	1.375	0.167
4차 조사: 지지후보 밝힘	0.000		1.000	
DK 응답	-0.504	0.338	0.604	0.137
조사불가	-3.545	0.161	0.029	0.000

V. 패널 이탈과 조건화가 최종 지지후보 및 투표의향에 영향을 주는가?

3절과 4절에 걸쳐 패널조사에서 패널 이탈과 조건화가 전개되는 양태를 살펴보았다. 패널조사가 시점 간 응답자의 의견 변화를 탐색하는 가장 좋은 방법이긴 하지만, 응답자의 반복 측정에 의해 선거조사에서 가장 중요한 지지후보와 투표의향에 부적절한 영향을 준다면 선거 패널조사의 가치는 회의적일 수밖에 없다.

KBS-MBMR의 2007년 대통령 선거 패널조사는 매 라운드마다 추가패널을 확보하도록 설계되었다. 따라서 진입 라운드별로 최종 지지후보 분포가 얼마나 다른가를 살펴보는 것이 가능하다. <표 17>이 그 결과인데 진입 라운드에 따라 최종 5차 조사에서의 지지후보 분포가 통계적으로 유의하게 다른 것으로 나타났다[카이제곱 = 52.9 (자유도 16), p -값 0.000]. 그러나 <표 17>에서 눈에 띄는 것은 없다/모른다/무응답(DK) 비율이 진입 라운드별로 큰 차이를 보인다는 점이다. 즉 1차 라운드 확보 패널에서는 DK 응답 비율이 6.1%로 낮았고 5차 라운드 확보 패널에서는 DK 응답 비율이 16.0%로 컸다. 3절에서 살펴본 바와 같이, 조기 라운드 확보 패널의 정치적 무관심자들은 후속 라운드에서 이탈하거나 지지후보 응답으로 태도를 바꾸는 패널 조건화가 되기 때문에 최종 조사에서는 DK 응답률이 작아진 것으로 보인다.⁵⁾ 이런 이유로 인하여 최종 지지후보 응답에서 DK 응답 비율이 진입 라운드별로 차이를 나타낸 것으로 생각된다.

DK 응답을 제외한 지지후보 분포를 진입 라운드별로 비교한 <표 18>을 보면, 진입 라운드 간의 순수 지지후보의 분포 차이는 통계적으로 유의하지 않은 것으로 나타났다 (카이제곱 = 11.9 (자유도 12), p -값 0.454). 이는 선거조사 패널이 이탈과 조건화로 영향을 받았음에도 불구하고 지지후보 분포가 일방적인 영향을 받지 않는 것을 의미하는 것이다. 마지막으로 <표 19>는 주민등록 상주인구 기준 지역·성·연령대 가중치를 적용하여 구한 KBS-MBMR 패널에서 DK 응답을 제외한 최종 지지후보 분포와 실제 선거결과를 비교한 것이다. 이명박 후보에 대한 지지율이 패널에서 49.6%, 선거에서 48.7%로 나타나 차이가 불과 0.9%P였다. 정동영 후보에 대하여는 패널에서 22.3%, 선거에서 26.1%로 나타나 차이가 -3.8%P로 다소 크게 나왔지만 이회창 후보에 대하여는 패널에서 15.2%, 선거에서 15.1%로 차이가 거의 없었다.

5) 선거를 1주일 정도 앞서 단면적 조사로 실시된 KBS-MBS 컨소시엄 선거예측 조사에서 무응답 비율은 RDD 표본(크기 1,533)에서 18.1%, 전화번호부 표본(크기 1,503)에서 15.2%를 기록하였다(허명희·김영원 2008). 패널 조건화가 누적된 1차 라운드 확보 패널의 DK 응답률은 6.1%로 이보다 훨씬 작다.

다음으로 진입 라운드별로 최종 5차 조사에서의 투표의향을 살펴보기로 한다. 패널 가입기간이 길어질수록 투표의향이 높아질까? <표 20>에 의하면 1차 라운드 확보 패널에서는 “반드시 투표할 것이다”가 76.2%로, 이후 패널의 78.4%, 81.0%, 89.5%, 79.9% 등에 비해 오히려 작았다. 선거에 관심이 적은 패널원들이 이탈해 갔고, 남은 패널원들은 반복 질문으로 인하여 선거에 관심이 커져 오래된 패널일수록 투표의향이 높을 것으로 기대되지만 실제로는 그렇지 않은 것이다. 따라서 패널조사가 인위적으로 투표의향을 높인다고 말하기 어렵다.

투표가 사회적으로 바람직한 행위이므로 조사면접에서 응답자들은 실제 생각보다 투표할 생각이 더 있는 것처럼 반응한다는 사회적 요망성 편향(social desirability bias) 때문에 투표율 예측값은 과다하게 잡히게 된다. 이를 바로 잡기 위하여 투표율 예측 시 미국 갤럽과 같이 투표의향을 수량화할 필요가 있다(Voss et al. 1995). 우리나라 실정에 맞는 수량화를 도출하기 위해서는 지속적인 시행오차(trial-and-error)가 불가피할 것이므로 다소 막연하더라도 첫 시도가 필요하다. 그 일환으로, “반드시 하겠다”에 0.8을, “웬만하면 할 것이다”에 0.4를, “반반이다”에 0.2를, “별로 생각이 없다”에 0.1을, “전혀 생각이 없다”에 0.05를 주고, 최종 5차 조사에서 조사가 되지 않은 나머지에 “반반이다”의 0.2를 부여해 보았다. 그 결과가 <표 21>인데 투표율 예측값이 61.5%로 나왔으며, 최종조사 응답자 표본에 지역·성·연령대에 가중치를 부여하는 경우 이 수치는 61.1%가 되었다. 2007년 12월 대선에서의 실제 투표율은 62.9%였다.

<표 17> 진입 라운드별 최종 조사에서의 지지후보 분포: DK 응답 포함

	표본 수	정동영	이명박	이회창	기타후보	DK	합계
1차 라운드	1,650	22.2%	49.6%	14.6%	12.1%	6.1%	100.0%
2차	361	17.5%	48.8%	15.2%	9.7%	8.9%	100.0%
3차	179	26.3%	45.8%	13.4%	11.2%	3.4%	100.0%
4차	171	15.2%	43.9%	17.0%	10.5%	13.5%	100.0%
5차	219	16.0%	45.2%	12.8%	10.0%	16.0%	100.0%
전체	2,580	20.9%	46.6%	14.0%	11.0%	7.6%	100.0%

* Pearson 카이제곱 = 52.9(자유도 16), p-값 0.000

<표 18> 진입 라운드별 최종 조사에서의 지지후보 분포: DK 응답 제외

(단위: 명)

	표본 수	정동영	이명박	이회창	기타후보	합계
1차 라운드	1,550	23.7%	49.6%	14.6%	12.1%	100.0%
2차	329	19.1%	53.5%	16.7%	10.6%	100.0%
3차	173	27.2%	47.4%	13.9%	11.6%	100.0%
4차	148	17.6%	50.7%	19.6%	12.2%	100.0%
5차	184	19.0%	53.8%	15.2%	12.0%	100.0%
전체	2,384	22.6%	50.4%	15.2%	11.9%	100.0%

* Pearson 카이제곱 = 11.9(자유도 12), p-값 0.454

<표 19> 가중치 적용 KBS-MBMR 패널의 최종 지지후보 분포와 선거 결과의 대비

	정동영	이명박	이회창	기타후보	합계
KBS-MBMR 패널	22.3%	49.6%	15.2%	12.9%	100.0%
선거결과	26.1%	48.7%	15.1%	10.2%	100.0%

<표 20> 진입 라운드별 최종 조사에서의 투표의향

(단위: 명)

	표본 수	반드시	할 것	반반	안할 것	전혀	합계
1차 라운드	1,650	76.2%	14.2%	7.3%	1.6%	0.7%	100.0%
2차	361	78.4%	15.8%	3.3%	2.2%	0.3%	100.0%
3차	179	81.0%	14.5%	1.7%	2.2%	0.6%	100.0%
4차	171	89.5%	5.3%	4.1%	0.6%	0.6%	100.0%
5차	219	79.9%	12.8%	4.6%	0.9%	1.8%	100.0%
전체	2,580	78.1%	13.7%	5.9%	1.6%	0.7%	100.0%

* Pearson 카이제곱 = 38.5(자유도 16), p-값 0.001.

* 선형적 연관성 카이제곱 = 6.8(자유도 1), p-값 0.009

〈표 21〉 투표의향의 수량화에 의한 투표율 예측

(단위: 명)

	표본 수	투표의향					조사불가
		반드시	할 것	반반	안할 것	전혀	
전체 패널	3,065	2,014	354	152	41	19	485
평균	61.5%	80%	40%	20%	10%	5%	20%
전체 패널 (가중치)	3,065	1,986	373	155	44	22	485
평균 (가중치)	61.1%	80%	40%	20%	10%	5%	20%

VI. 맺음 말

패널 이탈과 조건화는 패널 편향을 초래할 가능성이 있는 2개의 큰 요인이다. 그러나 패널 이탈과 조건화가 반드시 패널조사의 타당성을 저해하지는 않는다는 연구사례들이 보고되고 있다. 미디어 소비와 인터넷 활용을 조사한 Knowledge Networks Panel 연구가 그런 예의 하나이다(Clinton 2001). 본 소고의 결론은 KBS-MBMR의 2007년 대통령 선거 패널조사도 그렇다는 것이다. 즉, 여성과 20대·60대 및 없다/모른다/무응답(DK) 응답자들이 패널에서 이탈하는 성향이 있었고 DK 응답자들이 선호후보 밝힘 응답으로 급속히 전환되는 경향이 있었으나, 그럼에도 불구하고 선거 패널조사가 최종적으로 선호후보 밝힘 응답자 중 지지율 분포에 영향을 주었거나 투표의향을 인위적으로 높이지는 않았음을 보고하는 것으로 본 연구를 매듭짓는다.

참고문헌

- 강현철·한상태·김지연·정용찬·허명희, 2008, “RDD 전화조사와 주요 결과,” 《조사연구》 9(1): 1-22.
- 김장수, 2007, “선거연구와 패널여론조사: 2006년 지방선거를 중심으로,” 《조사연구》 8(1): 81-104.
- 이내영·이현우·김장수, 2007, 《변화하는 한국유권자: 패널조사를 통해 본 5·31 지방선거》, 동아시아연구원(EAI).
- 이현우·권혁용, 2008, 《변화하는 한국유권자 2: 패널조사를 통해 본 2007 대선》, 동아시아연구원(EAI).

- 허명희·김영원. 2008. "RDD 표본 대 전화번호부 표본: 2007년 대통령 선거 예측 사례," 《조사연구》 9(3): 55-69.
- 허명희·황진모. 2006. "전화조사를 위한 시간균형할당표본추출," 《조사연구》 7(2): 39-52.
- Bartels, L.M. 1999. "Panel Effects in the American National Election Studies," *Political Analysis* 8: 1-20.
- Clinton, J.D. 2001. "Panel Bias from Attrition and Conditioning: A Case Study of the Knowledge Networks Panel," presented at the 2001 AAPOR Conference in Montreal, Canada.
- Das, M., Toepoel, V. and Van Soest, A. 2007. "Can I Use a Panel? Panel Conditioning and Attrition Bias in Panel Surveys," *Center Discussion Paper Series No.* 2007-56.
- Dennis, M. and Li, R. 2003. "Effects of Panel Attrition on Survey Estimates," presented at the 2003 Annual Meeting of the American Association for Public Opinion Research in Nashville, TN.
- Loosveldt, G. and Carton, A. 1997. "Evaluation of Nonresponse in the Belgian Election Panel Study '91-'95." *Proceedings of the Survey Research Methods Section.* American Statistical Association.
- Sturgis, P., Allum, N. and Brunton-Smith, I. 2007. "Attitudes overTime: The Psychology of Panel Conditioning," in P. Lynn (ed.), *Methodology in Longitudinal Surveys* 1-13. Wiley.
- Taylor, B., Health A. and Lynn, P. 1996. "The Nature of Attrition in the British Election Panel Study," presented at the Fourth International Social Science Methodology Conference, Essex 1996.
- Voss, D., Gelman, A., and King, G. 1995. "Preelection Survey Methodology: Details from Eight Polling Organizations, 1988 and 1992." *Public Opinion Quarterly* 59: 98-132.
- Zargosky, J. and Rhoton, P. 1999. "Attrition and the National Longitudinal Survey's Women Cohorts." *Manuscript, Center for Human Resource Research.* Ohio State University.

[접수 2008/11/10, 수정 2009/2/20, 게재확정 2009/3/9]