

연구노트

2016년 국회의원선거 전화여론조사 정확성 분석

Analysis of Accuracy of Telephone Survey for the 2016 National Assembly Elections

김영원^{a)} · 윤지혜^{b)}

Young-won Kim · Jihye Yun

유권자들의 올바른 투표행위를 돕기 위해서는 선거여론조사를 통해 객관적이고 신뢰성 있는 정보를 제공해야 한다. 따라서 정당, 언론기관, 조사기관 등 여론조사 관련 기관에서는 여론조사의 정확성 제고를 위한 끊임없는 노력을 해왔다. 본 논문에서는 2016년 20대 국회의원선거 전화여론조사를 대상으로 하여 여론조사의 정확성에 관한 사후분석을 실시하였다. Martin et al.(2005)이 제시한 정확성 척도 A 를 기준으로 예측오차와 편향을 측정하여 조사방법이나 응답률, 가중치 적용 등이 선거여론조사의 정확성에 미치는 영향을 분석하였다. 분석의 결과 ARS나 ARS와 전화면접을 혼합한 방식보다 전화면접 방식을 통해 보다 정확한 여론조사 결과를 얻을 수 있는 것으로 나타났다. 또한 응답률을 높이기 위한 노력과 연령대·성별 할당을 채우는 노력 등을 통해 여론조사의 정확성을 높일 수 있다는 점도 확인할 수 있었다.

주제어 : 선거여론조사, 전화조사, 예측오차, 편향, 20대 국회의원선거

We investigate the forecast error and bias of public opinion polls conducted for 20th national assembly elections in Korea. The

a) 교신저자(corresponding author): 숙명여자대학교 통계학과 교수 김영원.

E-mail: ywkim@sm.ac.kr

b) 숙명여자대학교 통계학과 대학원생.

aim of this study is to find main cause of the forecasting error and bias of telephone survey. We use the telephone survey paradata released by National Fair Election Survey Deliberation Commission and the measure of predictive accuracy(statistic A) proposed by Martin et al.(2005). By analyzing the 203 public opinion polls carried out in Seoul, Incheon and Gyeonggi province, we found that telephone survey method, period of poll, variation level of weights, and response rate are the factors to cause the forecasting error and bias in telephone survey.

Key words: public opinion poll, telephone survey, forecasting error, bias, 20th national assembly election

I. 서론

선거여론조사는 단순히 여론의 동향을 파악하는 데 그치는 것이 아니라 여론의 형성 및 변화에도 영향을 미친다는 점에서 선거과정에 있어 매우 중요한 역할을 한다. 특히 2016년 20대 국회의원선거에서는 여론조사를 통한 당내 경선이 실시되어 여론조사의 중요성이 한층 더 높아졌다. 그러나 여론조사의 중요성과 함께 여론조사에 의한 여론 왜곡 및 조작에 대한 논란이 끊이질 않았고, 정확성에 대한 의구심 또한 높아졌기 때문에 여론조사의 객관성과 신뢰성 제고를 위해서는 여론조사의 정확성에 대한 과학적이고 실증적인 연구가 필요하다.

여론조사의 정확성을 판단하기 위해서는 당선자를 옳게 예측했는지의 여부뿐만 아니라, 각 후보자의 득표율을 얼마나 정확하게 추정했는지도 고려해야 한다. 따라서 예측오차를 기반으로 한 통계학적인 관점에서 여론조사의 정확성 평가가 필요하며, 예측오차와 함께 오차의 방향과 편향의 발생 여부에 대한 연구가 함께 실시되어야 한다.

여론조사의 정확성을 평가하는 예측정확도 측정을 위해 Martin et al.(2005)은 정확성 척도 A 를 제시하였다. 척도 A 는 예측오차와 편향을 측정할 수 있으며, 후보자의 수와 부동층의 처리방법에 영향을 받지 않는 등 많은 장점을 지니고 있다. 따라서 본 연구에서는 Martin et al.(2005)이 제시한 정확성 척도 A 를 기반으로 20대 국회의원선거 여론조사의 정확성에 대한 연구를 실시한다.

참고로 미국과 영국에서도 최근 선거여론조사에서 오차의 발생 요인을 규명하기 위한 연구들이 진행된 바 있다. 미국의 경우 2008년 미국의 각 정당별 대통령 예비선거에 대한 여론조사의 정확성에 대해 많은 비판이 있었다. 이에 따라 AAPOR 주관으로 Bolger et al.(2009)이 수행한 연구에서는 어떤 전화조사 방식을 적용했는지, 무선전화만을 사용하는 사람들이 추출틀에 포함되었는지, 무응답으로 인해 특정 계층이 조사에 적게 포함되었는지, 질문방식과 후보자 나열 순서 등 설문지 구성, 응답자 중 투표자를 구분하기 위해 사용한 선별 문항의 적절성과 투표가능성 가중치 계산방식의 정확성, 전화 접촉에 실패한 경우 응답을 얻기 위해 시도한 전화 재접촉 횟수, 유권자들이 지지 후보를 결정하는 시점과 관련된 문제, 응답자들이 사회적으로 바람직한 응답을 하려는 경향에 따른 편향, 모집단 특성에 대한 벤치마킹 과정에서 발생하는 잘못된 가중치 부여, 부동층 할당 과정에서 발생하는 오류 등에 따른 오차의 발생 가능성을 종합적으로 분석하였다.

한편 영국에서는 2015년 총선 여론조사에서 발생한 보수당 의석 수 과소 예측에 대한 심각한 문제 제기가 있었으며, 이에 따라 BPC(British Polling Council)와 MRS(Market Research Society)는 2015년 영국 총선 여론조사의 예측오차 원인을 규명하기 위한 연구를 수행하였다(Sturgis et al. 2016). Sturgis et al. (2016)이 수행한 연구에서는 2015년 영국 총선 여론조사 예측오차 발생 원인으로 다음 몇 가지 사항을 검토하였다. 첫째, 부재자투표자와 해외투표자 처리와 함께 여론조사 응답자 중 실제로 투표를 하지 않을 유권자를 구분해 처리하는 문제, 둘째, 질문의 배열과 형식에 따라 응답이 달라질 수 있다는 점, 셋째, 투표일에 임박하여 부동층이 지지 정당을 결정하는 경우를 포함해 막판에 유권자들이 지지 정당을 변경함으로써 발생하는 문제, 넷째, ‘수줍은 보수당’ 효과 등 사회적으로 바람직한 응답을 하려는 경향에 따른 편향, 다섯째, 응답자가 실제로 투표할

가능성이 어느 정도인지를 추정에 반영하기 위해 사용하는 투표가능성 가중치 부여 방법, 여섯째, 조사방법에 따른 효과, 마지막으로, 엄격한 확률표본추출이 이루어지지 못함으로써 발생하는 표본의 대표성 결여와 관련된 문제를 지적하고 있으며, 특히 표본의 대표성이 결여가 예측오차에 가장 큰 영향을 주는 요인이라고 결론내리고 있다.

본 연구에서는 미국이나 영국의 사례에서와 같이 대규모 연구를 통해 우리나라 여론조사의 정확성에 미치는 영향을 심층적으로 분석하지는 못하였지만 주어진 자료를 최대한 활용하여 관련 연구를 수행하였다. 2장에서는 여론조사 정확도 척도와 분석대상 여론조사에 대해 간략하게 소개한다. 3장에서는 2016년 20대 국회의원선거 여론조사를 대상으로 지역, 조사기관, 조사방법, 조사시점, 응답률 및 가중치 변동 등에 따른 여론조사의 정확성을 비교함으로써 여론조사의 정확성에 영향을 미치는 요인들을 살펴본다. 마지막으로 4장에서는 전국 정당지지율에 관한 여론조사를 대상으로 시간의 흐름에 따른 정확성의 변화 추이를 살펴보고자 한다.

II. 여론조사 정확성 척도와 분석대상 여론조사

전화여론조사를 구체적으로 수행하는 방법에 따라 조사의 정확성이나 편향의 발생 여부 등에 차이가 있기 때문에 이를 평가하기 위해 본 연구에서는 Martin et al.(2005)이 제시한 정확성 척도 A 를 사용하기로 한다. 김영원·최윤정(2011)과 김영원·황다솜(2014)은 우리나라 선거여론조사의 정확성 분석을 위해 척도 A 를 사용하였으며, 최근 국내·외 여론조사의 오차분석을 위한 연구에서도 척도 A 가 폭넓게 사용되고 있다(최필선·민인식 2013; Mokrzycki et al. 2009; Traugott & Christopher 2009; Bolger et al. 2009).

참고로 척도 A 는 계산이 쉽고 비교가 용이하다는 것 외에도 몇 가지 장점을 가진다. 첫 번째는 표준화된 값이기 때문에 다른 선거의 여론조사와도 비교가 가능하다. 두 번째는 그룹으로 묶어서 비교할 수 있다. 예를 들어, 두 정당만을

비교하는 것이 아니라 정권을 잡지 않은 정당들을 야당으로 묶어 여당과 야당으로도 비교가 가능하다. 세 번째는 편향이 유의하게 발생하였는지의 여부를 판단할 수 있다. 네 번째는 후보자의 수가 달라도 비교가 용이하며, 다섯 번째는 부동층의 처리 방법에 따른 영향을 적게 받는다. 또한 다변량분석이나 회귀분석 등을 통해 정확성에 영향을 주는 요인들을 파악할 수 있다. 마지막으로 시간에 따라 A 값이 변화하는 추이를 분석할 수도 있다.

본 연구에서 s, t, p 는 각각 ‘새누리당’, ‘더불어민주당’(이하 ‘더민주당’), ‘국민의당’의 후보자 예측 득표율 혹은 정당지지율이고, S, T, P 는 개표결과 새누리당, 더민주당, 국민의당 후보자의 실제 득표율 혹은 정당지지율을 나타낸다. 새누리당과 더민주당의 자료를 가지고 계산한 A 는 다음과 같다.

$$A = \log\left[\frac{s/t}{S/T}\right] \quad (1)$$

A 가 0이면 새누리당과 더민주당의 여론조사 결과가 선거 결과와 일치하는 것이고, 0보다 크면(작으면) 새누리당(더민주당)이 여론조사에서 과대추정되었다는 것을 의미한다.

여기서는 척도 A 를 기준으로 조사방법이나 가중치 및 응답률 차이 등에 따른 여론조사의 정확성을 비교하고자 한다. 개별 여론조사(i)의 예측정확도 A_i 를 계산하고, 이를 토대로 분석대상에 따라 해당 범주에 속하는 A_i 들의 평균($\bar{A}$)을 사용하여 예측오차의 방향과 편향을 살펴본다. 한편, A_i 를 합치는 과정에서 부호에 따라 서로 상쇄될 수 있다는 점을 감안해 평균적인 오차를 보기 위해 A_i 들의 절댓값 평균($\overline{|A|}$)을 함께 살펴본다. 아울러 A_i 들의 표준편차($s.d.(A)$)를 구해 95% 신뢰구간을 계산하고, 이를 기준으로 신뢰구간을 벗어나는 경우 표본추출 오차로는 설명할 수 없는 수준의 심각한 치우침이 발생한 것으로 보고 이를 편향이 발생한 것으로 간주한다. n 은 분석대상 범주에 속한 전체 여론조사 수를 나타내고, $n^+(n^-)$ 는 이 중 양(음)의 방향으로 발행한 편향된 여론조사의 수를 나타낸다.

본 연구에서는 2016년 4월 13일에 치러진 20대 국회의원선거 여론조사를 대상으로 정확성을 분석하기로 한다. 중앙선거여론조사공정심의위원회(이하 ‘여심

위')에 등록되어 있는 20대 국회의원선거 관련 여론조사는 총 1,410개이며, 이중 일정 조건에 부합하는 여론조사들을 대상으로 분석하기로 한다.

오차 분석을 위해 여론조사 결과와 실제 득표율을 비교해야 하므로 후보자 등록이 완료된 3월 26일부터 여론조사 공표가 가능한 4월 6일까지 실시된 여론조사를 분석하였다. 지역적으로 보면 전국의 모든 선거구에서 이루어진 여론조사를 분석할 수도 있지만 여기서는 서울, 경기, 인천을 포함하는 수도권 선거구 중 1, 2위가 모두 새누리당과 더민주당 후보가 차지한 79개 선거구의 203개 여론조사를 분석대상으로 하였다.

실제로 전국 선거구에서 수행된 여론조사를 대상으로 하는 경우에도 본 연구에서 제시하는 분석결과와 거의 유사하다. 참고로 윤지혜(2016)는 본 연구에 제시한 결과와 함께 전국의 선거구 중 1위 또는 2위 후보자가 새누리당과 더민주당 후보자이면서 4회 이상 여론조사가 실시된 선거구를 대상으로 한 분석결과를 제시하고 있으며, 분석결과를 보면 본 연구결과와 큰 차이가 없음을 확인할 수 있다.

Ⅲ. 여론조사 정확성 분석결과

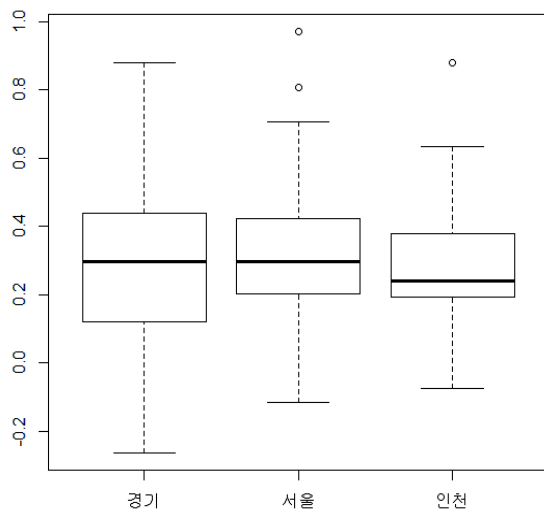
1. 지역별 예측오차 비교

수도권에 해당하는 경기, 서울, 인천 지역 선거구에서 수행된 여론조사의 지역별 편향 및 예측오차 분석결과는 <표 1>과 같고, A_i 와 $|A_i|$ 의 상자그림은 <그림 1>과 <그림 2>에 제시되어 있다. 203개의 여론조사 중 경기 지역에 108개(53.20%), 서울 지역에 76개(37.44%), 인천 지역에 19개(9.36%)의 여론조사가 분석대상이었으며, 경기 지역에서 실시된 여론조사 108개 중 46개(42.59%) 여론조사에서 편향이 발생하였다. 서울 지역에서는 76개의 여론조사 중 36개(47.37%) 여론조사에서 편향이 발생하였고, 인천 지역에서는 19개의 여론조사 중 6개(31.58%)의 여론조사에서 편향이 발생하였다.

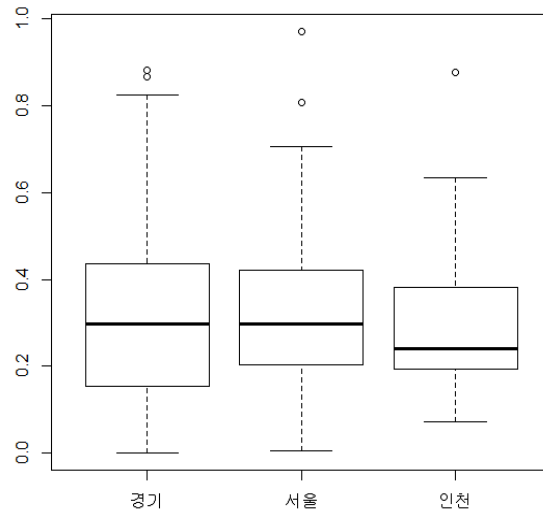
새누리당과 더불어민주당이 1, 2위를 차지한 20대 국회의원 수도권 여론조사에서는 표본추출오차 수준을 넘어서는 편향이 발생한 경우 모두 새누리당을 과대추정하는 양의 편향이 발생했다는 것을 확인할 수 있다. $\bar{A}$ 와 $|\bar{A}|$ 이 서울, 인천, 경기 순으로 크게 나타났다. $\bar{A}$ 는 수도권 모든 지역에서 양수 값을 가지며, 경기 지역이 가장 0에 가까웠다. 즉, 수도권 지역 중 경기 지역이 예측오차도 가장 작고, 새누리당에 대한 치우침도 가장 작았다.

<표 1> 지역별 편향 및 예측오차

지역	$\bar{A}$	$ \bar{A} $	$s.d.(A)$	n	n^+	n^-	편향 비율
경기	0.2842	0.3080	0.2376	108	46	0	0.4259
서울	0.3036	0.3160	0.2037	76	36	0	0.4737
인천	0.3015	0.3093	0.2182	19	6	0	0.3158



<그림 1> 지역별, A_i 의 분포



<그림 2> 지역별 $|A_i|$ 의 분포

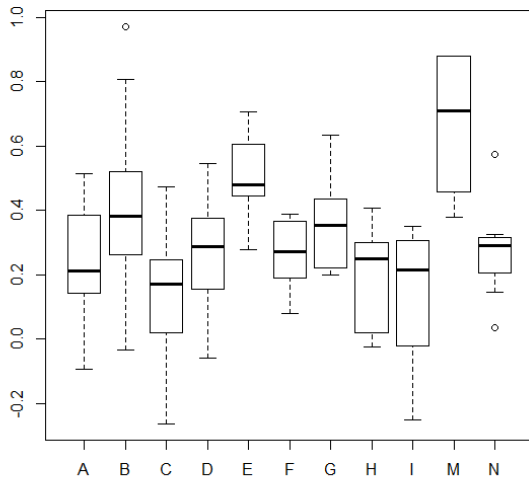
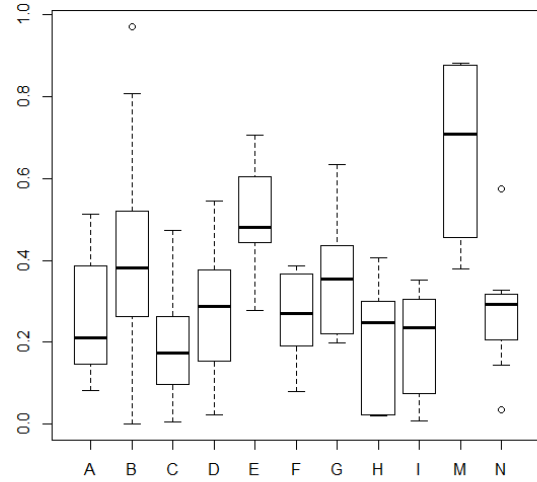
2. 조사기관별 예측오차 비교

수도권 203개의 여론조사는 총 23개 조사기관이 수행하였다. 이 중 5번 이상의 여론조사를 실시한 조사기관에 대해서 조사기관별 편향 및 예측오차를 분석하였다. 기관별 조사 횟수는 5~55번인 것으로 나타났고, 조사기관 예측오차 비교 분석에 사용된 전체 여론조사의 수는 184개이다.

<표 2>에 조사기관별 편향 및 예측오차가 정리되어 있고, <그림 3>과 <그림 4>에 A_i 와 $|A_i|$ 의 상자그림이 제시되어 있다. 모든 조사기관의 $\bar{A}$ 가 양수로 나타났다. M사와 E사는 $\bar{A}$ 와 $|\bar{A}|$ 모두 가장 크게 나타났다. 다시 말해, M사와 E사는 예측오차도 크고 새누리당에 대한 치우침도 가장 크게 발생했다. 반면에 I사와 C사는 $\bar{A}$ 와 $|\bar{A}|$ 모두 상대적으로 가장 작은 것으로 나타났으며, 결과적으로 다른 조사기관에 비해 정확성이 우수했던 것으로 볼 수 있다.

<표 2> 조사기관별 편향 및 예측오차

조사기관	$\bar{A}$	$ \bar{A} $	$s.d.(A)$	n	n^+	n^-	편향 비율
A	0.2410	0.2677	0.2069	7	3	0	0.4286
B	0.3844	0.3857	0.1974	55	36	0	0.6545
C	0.1387	0.1870	0.1668	49	5	0	0.1020
D	0.2536	0.2642	0.1758	15	5	0	0.3333
E	0.5023	0.5023	0.1631	5	4	0	0.8000
F	0.2596	0.2596	0.1275	5	2	0	0.4000
G	0.3568	0.3568	0.1607	7	4	0	0.5714
H	0.2001	0.2077	0.1664	6	2	0	0.3333
I	0.1333	0.1936	0.1944	13	1	0	0.0769
M	0.6689	0.6689	0.2223	6	6	0	1.0000
N	0.2781	0.2781	0.1673	7	1	0	0.1429

〈그림 3〉 조사기관별 A_i 의 분포〈그림 4〉 조사기관별 $|A_i|$ 의 분포<표 3> 조사기관별 $|A_i|$ 의 LSD 다중비교 결과

조사기관	평균	그룹
M	0.6689	a
E	0.5023	ab
B	0.3857	bc
G	0.3568	bcd
N	0.2781	cde
A	0.2677	cde
D	0.2642	de
F	0.2596	de
H	0.2077	de
I	0.1936	e
C	0.1870	e

조사기관별 예측오차가 통계적으로 유의한 차이가 있는지 보기 위해, 오차를 의미하는 $|A_i|$ 들을 기준으로 분산분석을 수행하였다. F 값이 9.049이고 p 값이

0.0001보다 작아 통계적으로 유의한 차이가 존재했다. 조사기관별 $|A_i|$ 들의 차이에 대해 자세히 알아보기 위해 Fisher의 LSD 다중비교를 실시하였다. 다중비교 결과는 가독성을 높이기 위해 $|A_i|$ 들을 기준으로 유의한 차이가 있는지 여부를 볼 수 있는 그룹표 형식으로 <표 3>에 정리되어 있다. 예측오차를 의미하는 $|A_i|$ 들을 기준으로 조사기관들을 나누면, M사와 E사가 $|A_i|$ 들이 가장 큰 a그룹에 속하며 C사와 I사가 $|A_i|$ 들이 가장 작은 e그룹에 속한다. 따라서 M사와 E사는 상대적으로 다른 조사기관에 비해 예측오차가 컸고, C사와 I사는 상대적으로 다른 조사기관에 비해 예측오차가 작았다고 볼 수 있다.

3. 조사방법별 예측오차 비교

분석대상 203개 여론조사는 여심위 조사방법 분류방식에 의하면 전화면접, ARS 외에도 두 가지 이상의 방법이 혼합된 매우 다양한 방식으로 구분될 수 있다. 여기서는 의미 있는 조사방법에 따른 정확성 비교를 위해 조사방법을 네 가지로 재분류하였다.

먼저, 여론조사의 80% 이상을 전화면접 또는 ARS 조사로 수행된 경우가 있었으며, 이 경우 해당 조사는 전화면접 또는 ARS 조사로 분류하였다. 전화면접과 ARS 두 가지 방식을 비슷한 비율로 혼합한 경우(전화면접 비율이 36%~41%, ARS 비율이 59%~64%)는 ARS&전화면접으로 분류하였다. 마지막으로 이런 세 가지 구분에 해당하지 않는 나머지 여론조사들은 전화면접과 ARS에 스마트폰앱조사와 인터넷조사 등을 다양한 비율로 혼합한 여론조사들이었다. 이런 기준으로 구분하면 전체 분석대상 여론조사는 조사방법에 따라 전화면접, ARS, ARS&전화면접, 혼합 등 네 가지 범주로 구분될 수 있다.

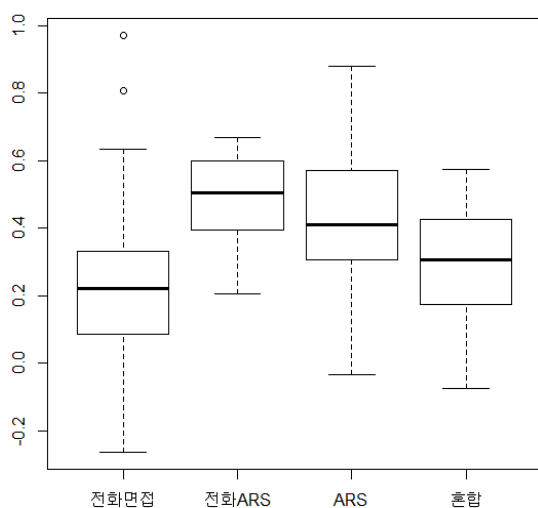
앞에서 설명한 분류기준에 따라 네 가지 범주의 조사방법의 예측오차와 편향에 대해 분석하였다. 조사방법별 예측오차와 편향에 대한 분석결과가 <표 4>에 정리되어 있고, <그림 5>와 <그림 6>에 A_i 와 $|A_i|$ 의 상자그림이 제시되어 있다.

분석대상 여론조사 중 전화면접으로 실시된 여론조사는 117개(57.64%)이고, ARS는 51개(25.12%), ARS&전화면접은 7개(3.45%), 혼합은 28개(13.79%)이

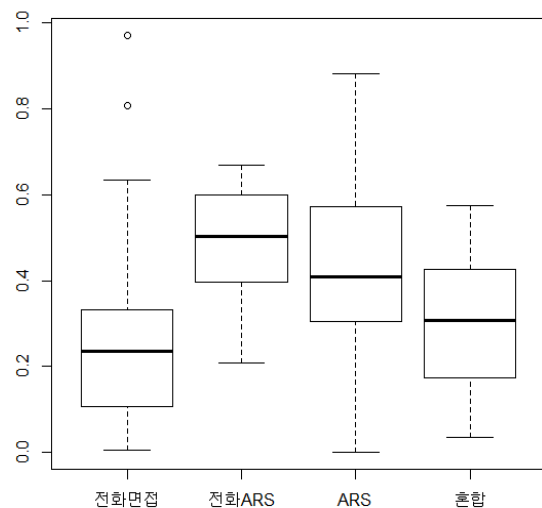
다. 전화면접으로 실시된 여론조사 117개 중 31개(26.50%) 여론조사에서 편향이 발생했으며, ARS는 51개의 여론조사 중 39개(76.47%), ARS&전화면접은 7개의 여론조사 중 6개(85.71%), 혼합은 28개의 여론조사 중 12개(42.86%) 여론조사에서 편향이 발생하였다. 편향비율은 전화면접이 가장 작았으며, 다른 조사방법의 경우 혼합, ARS, ARS&전화면접 순으로 편향비율이 작았다. $\bar{A}$ 와 $|A|$ 모두 전화면접에서 가장 작은 것으로 나타났고, ARS&전화면접의 경우 $\bar{A}$ 와 $|A|$ 모두 가장 큰 것으로 나타났다.

<표 4> 조사방법별 편향 및 예측오차

조사방법	$\bar{A}$	$ \bar{A} $	$s.d.(A)$	n	n^+	n^-	편향 비율
전화면접	0.2177	0.2471	0.2046	117	31	0	0.2650
ARS&전화면접	0.4819	0.4819	0.1661	7	6	0	0.8571
ARS	0.4350	0.4364	0.2161	51	39	0	0.7647
혼합	0.3024	0.3077	0.1641	28	12	0	0.4286



<그림 5> 조사방법별 A_i 의 분포



<그림 6> 조사방법별 $|A_i|$ 의 분포

조사방법별로 예측오차를 나타내는 $|A_i|$ 가 통계적으로 유의한 차이가 있는지 확인하기 위해 분산분석을 실시한 결과, F 값이 15.54이고 p 값이 0.0001보다 작아 유의한 차이는 있었다. 조사방법별로 통계적으로 유의한 차이가 있는지 보기 위해 $|A_i|$ 를 가지고 Fisher의 LSD 다중비교를 실시했다. 그 결과는 <표 5>에 정리되어 있으며, 전화면접과 혼합이 ARS와 ARS&전화면접보다 $|A_i|$ 가 작은 것으로 나타났다. 다시 말해, 전화면접과 혼합이 ARS와 ARS&전화면접보다 통계적으로 유의하게 예측오차가 더 작았다는 것을 알 수 있다.

<표 5> 조사방법별 $|A_i|$ 의 LSD 다중비교 결과(수도권)

조사방법 비교	차이	p 값	유의여부
ARS-ARS&전화면접	-0.0455	0.5274	
ARS-전화면접	0.1892	0.0000	***
ARS-혼합	0.1287	0.0025	**
ARS&전화면접-전화면접	0.2348	0.0009	***
ARS&전화면접-혼합	0.1742	0.0218	*
전화면접-혼합	-0.0605	0.1084	

참고로 앞의 2절 조사기관별 예측오차 비교에서 상대적으로 오차가 큰 것으로 나타난 M사와 E사의 경우 ARS전화조사 방식으로 조사를 수행한 조사기관들이다. 이 중 M사의 경우 유선ARS 방식으로만 조사를 수행하였고, E사의 경우는 유선ARS 80% 정도에 나머지를 스마트폰앱조사로 채우는 방식으로 조사를 수행하였다. 반면 상대적으로 정확한 조사를 수행한 것으로 나타난 C사와 M사의 경우는 전화면접조사 방식을 위주로 조사를 수행한 조사기관들이다. C사의 경우 유선전화면접과 무선전화면접을 개략적으로 8:2 수준으로 병합하는 방식을 사용하였고, I사의 경우는 유선전화면접과 스마트폰앱조사를 8:2 수준으로 병합하는 방식으로 조사를 수행하였다. 이런 결과를 통해서도 ARS전화조사를 위주로 여론조사를 수행하는 경우 전화면접조사 방식에 비해 정확성이 상당히 떨어진다는 것을 다시 한 번 확인할 수 있다.

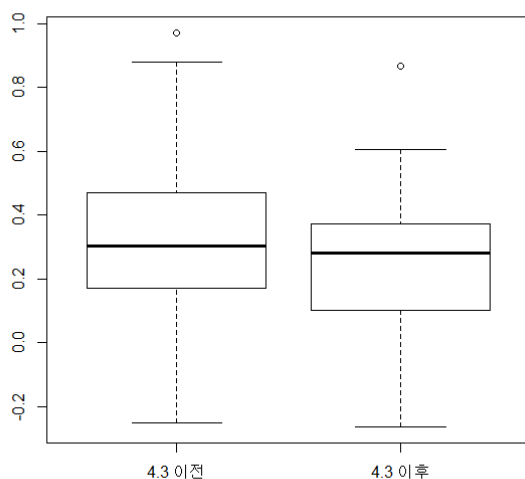
4. 조사시점에 따른 예측오차 비교

여론조사 시점에 부동층이었던 유권자들이 투표일에 임박해서 지지할 후보를 결정할 수도 있고 마지막 순간에 지지 후보자를 바꾸는 경우도 있기 때문에 조사 시점이 여론조사의 정확성에 영향을 줄 수 있다. 여기서는 조사시점에 따른 오차 비교를 위해 여론조사 공표금지 기간을 고려해 투표일 10일 전인 4월3일을 기준으로 이전과 이후에 실시된 조사를 비교하기로 한다.

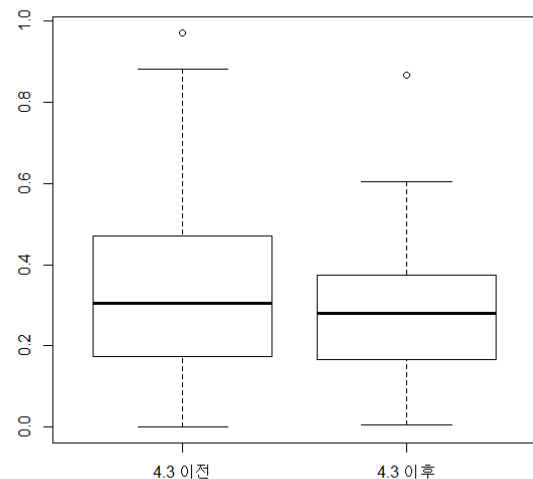
203개의 여론조사 중 4월3일 이전에 실시된 여론조사는 141개(69.46%)였고, 4월3일 이후에 실시된 여론조사는 62개(30.54%)였다. 편향과 예측오차는 <표 6>과 같고, A_i 와 $|A_i|$ 의 상자그림은 <그림 7>과 <그림 8>에 제시되어 있다. 4월3일 이전에 실시된 141개의 여론조사 중 61개(43.26%)의 여론조사에서 편향이 발생했으며, 4월3일 이후에 실시된 62개의 여론조사 중에서는 27개(43.55%)의 여론조사에서 편향이 발생하였다.

<표 6> 조사시점에 따른 편향 및 예측오차

조사시점	$\bar{A}$	$ \bar{A} $	$s.d.(A)$	n	n^+	n^-	편향비율
4.3 이전	0.3123	0.3280	0.2310	141	61	0	0.4326
4.3 이후	0.2495	0.2727	0.1977	62	27	0	0.4355



<그림 7> 조사시점에 따른 A_i 의 분포



<그림 8> 조사시점에 따른 $|A_i|$ 의 분포

오차를 의미하는 $\overline{|A|}$ 를 보면 4월 3일 이후보다 4월 3일 이전에 실시된 여론조사의 경우 더 크다는 것을 볼 수 있다. 조사시점에 따른 $|A_i|$ 들의 차이가 통계적으로 유의한지 보기 위해 t -test를 실시한 결과, t 값이 1.8546이고 p 값이 0.0651로 유의한 차이가 있다는 결론을 내릴 수 있었다. 즉, 20대 국회의원선거 수도권 여론조사에서 4월3일 이전의 여론조사에 비해 투표일이 가까운 시점에 실시된 4월3일 이후의 여론조사가 상대적으로 오차가 적었다고 말할 수 있다. 편향 또는 오차의 방향성을 나타내는 $\overline{A}$ 에 대해서도 유사한 해석이 가능하고, 4월3일 이전의 여론조사에 비해 4월3일 이후의 여론조사가 새누리당에 대한 치우침이 상대적으로 작았다고 말할 수 있다. 종합하면 투표일에 근접한 시점에 이루어진 조사가 상대적으로 정확성이 높은 것으로 결론내릴 수 있다.

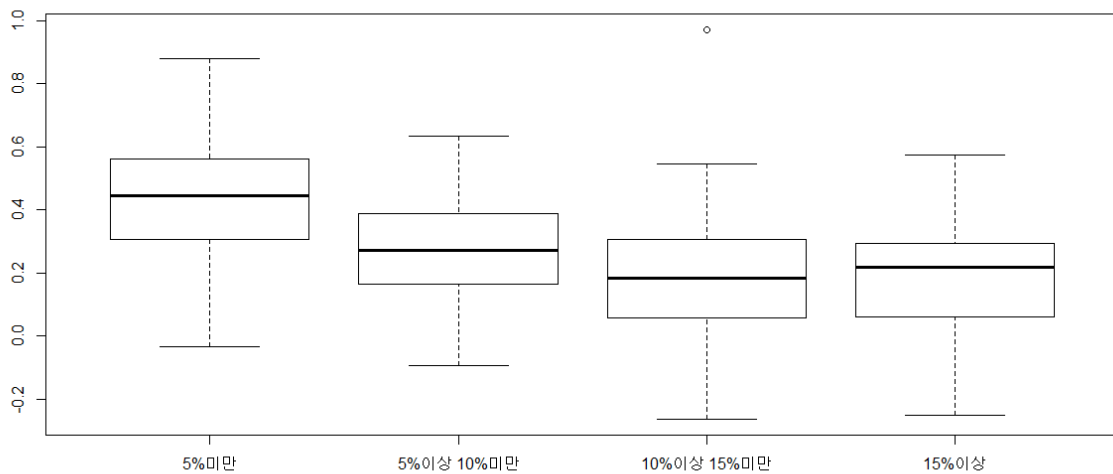
5. 응답률에 따른 예측오차 비교

일반적으로 응답률은 여론조사의 품질을 평가하는 중요한 요소라는 점을 고려해 여기서도 분석대상 여론조사의 응답률이 조사의 정확성에 미치는 영향을 살펴보고자 한다. 우선 응답률에 따른 편향과 예측오차는 <표 7>에 정리되어 있고, A_i 와 $|A_i|$ 의 상자그림이 <그림 9>와 <그림 10>에 제시되어 있다.

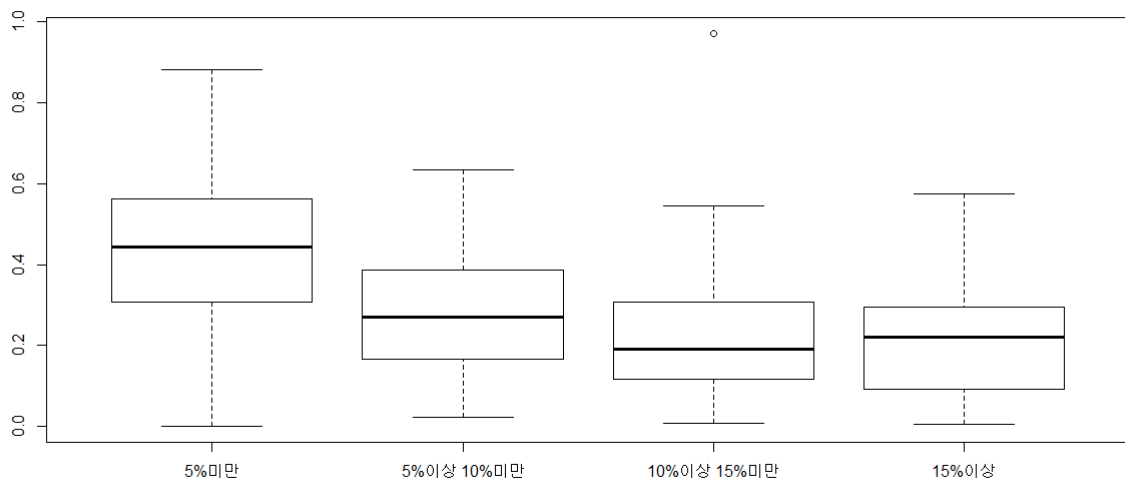
<표 7> 응답률에 따른 편향 및 예측오차

응답률	$\overline{A}$	$\overline{ A }$	$s.d.(A)$	n	n^+	n^-	편향 비율
5% 미만	0.4382	0.4392	0.1986	73	56	0	0.7671
5% 이상 10% 미만	0.2617	0.2737	0.1752	49	20	0	0.4082
10% 이상 15% 미만	0.1842	0.2303	0.2245	41	7	0	0.1707
15% 이상	0.1783	0.2061	0.1678	40	5	0	0.1250

<표 7>을 보면 응답률이 높을수록 편향비율이 감소하는 것을 확인할 수 있다. 또한 $\bar{A}$ 가 응답률이 높을수록 감소하는 것을 확인할 수 있다. 통계적으로 응답률에 따라 A_i 들이 유의하게 다른지 파악하기 위해 분산분석을 실시한 결과, F 값이 23.21이고 p 값이 0.0001보다 작게 나와 유의한 차이가 있다는 것을 알 수 있다.



〈그림 9〉 응답률에 따른 A_i 의 분포



〈그림 10〉 응답률에 따른 $|A_i|$ 의 분포

응답률별 자세한 차이를 알기 위해, A_i 를 가지고 Fisher의 LSD 다중비교를 실시하였다. 그 결과가 <표 8>에 제시되어 있다. 5% 미만인 여론조사가 5% 이상인 여론조사에 비해 A_i 가 크고, 5% 이상 10% 미만인 여론조사가 10% 이상인 여론조사에 비해 A_i 가 크다는 것을 알 수 있다. 따라서 5% 미만, 5% 이상 10% 미만, 10% 이상으로 응답률이 증가함에 따라 새누리당에 대한 치우침이 감소한다는 것을 확인할 수 있다. 모든 범주에서 $\bar{A}$ 가 0보다 크다는 점을 고려하면, 응답률이 높을수록 결과적으로 새누리당에 대한 치우침이 완화되어 정확성이 높아진다는 결론을 내릴 수 있다. $|\bar{A}|$ 에 대한 분석을 통해서도 유사한 결과를 얻을 수 있었으며, 자세한 분석결과는 생략하기로 한다.

<표 8> 응답률에 따른 A_i 의 LSD 다중비교 결과

응답률 비교	차이	p값	유의여부
5% 미만-5% 이상 10% 미만	0.1766	0.0000	***
5% 미만-10% 이상 15% 미만	0.2540	0.0000	***
5% 미만-15% 이상	0.2599	0.0000	***
5% 이상 10% 미만-10% 이상 15% 미만	0.0775	0.0596	
5% 이상 10% 미만-15% 이상	0.0834	0.0441	*
10% 이상 15% 미만-15% 이상	0.0059	0.8905	

한편, 응답률과 A_i , $|A_i|$ 는 모두 연속형 변수이기 때문에 상관계수를 계산하여 상관관계를 확인할 수 있다. 응답률과 A_i 의 상관계수를 구해보면 -0.4618이고, p값은 0.0001보다 작아 강한 음의 상관관계를 가진다는 것을 확인할 수 있다. 또한 A_i 의 부호가 거의 대부분 양수이기 때문에 응답률과 $|A_i|$ 의 상관계수도 -0.4600으로 거의 동일한 결과를 보여준다. 따라서 분석대상 여론조사의 경우, 응답률이 증가함에 따라 새누리당에 대한 치우침과 오차가 감소해 정확성이 높아진다는 결론을 내릴 수 있다.

6. 가중치 변동에 따른 예측오차 비교

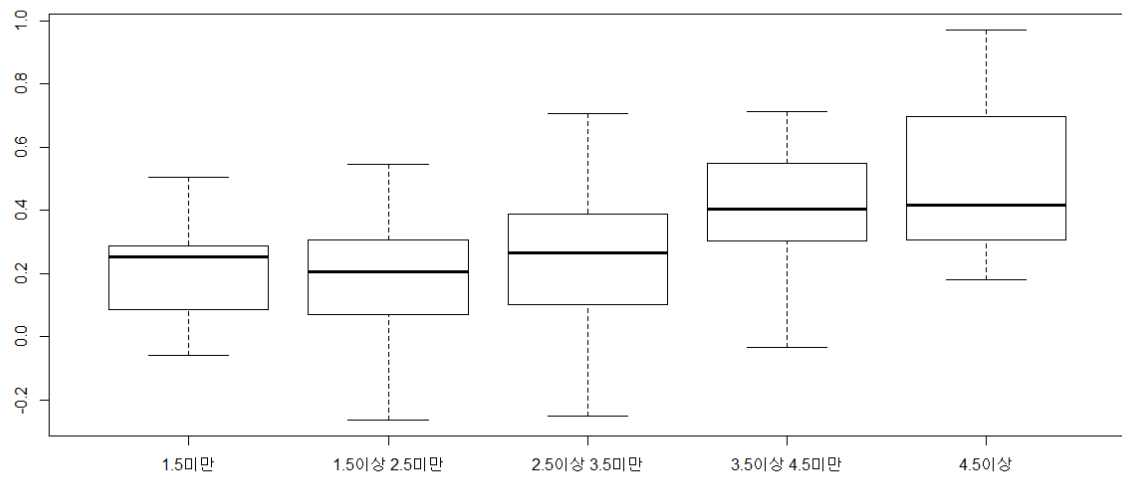
할당추출에 의한 여론조사에서 집단별 가중치가 변동이 크다는 것은 의도한 대로 할당을 제대로 채우지 못했다는 것을 의미하며, 이는 여론조사의 정확성을 떨어뜨리는 원인이 될 수 있다. 여심위 자료에서는 개별 가중치는 확인할 수 없고 성별 또는 연령대별 그룹별 가중치만을 확인할 수 있기 때문에 여기서는 연령대별 가중치 중 가장 큰 가중치를 가장 작은 가중치로 나눈 것을 가중치 비로 정의하고, 가중치 비를 가중치 변동을 나타내는 지표로 사용하였다. 가중치 비를 기준으로 여론조사들을 가중치 비가 1.5 미만, 1.5 이상 2.5 미만, 2.5 이상 3.5 미만, 3.5 이상 4.5 미만, 4.5 이상인 그룹으로 나누고, 그룹 간 비교를 통해 가중치 변동에 따른 여론조사의 정확성을 분석하였다.

가중치 비에 따른 예측오차와 편향에 대한 분석결과가 <표 9>에 정리되어 있고, <그림 11>과 <그림 12>에 A_i 와 $|A_i|$ 의 상자그림이 제시되어 있다.

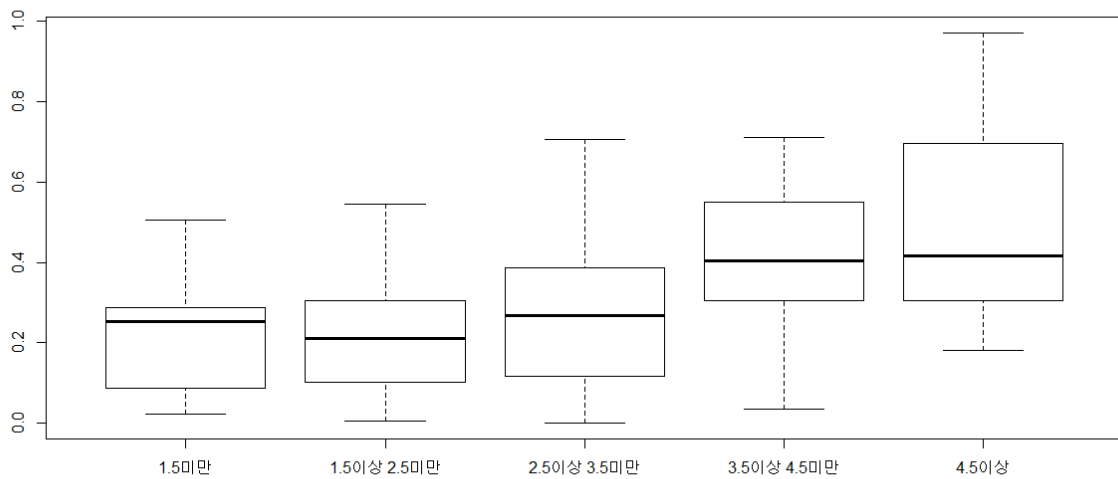
<표 9> 가중치 비에 따른 편향 및 예측오차

가중치 비	$\bar{A}$	$ \bar{A} $	$s.d.(A)$	n	n^+	n^-	편향 비율
1.5 미만	0.1994	0.2172	0.1758	9	1	0	0.1111
1.5 이상 2.5 미만	0.1798	0.2162	0.1778	70	13	0	0.1857
2.5 이상 3.5 미만	0.2529	0.2709	0.2074	49	18	0	0.3673
3.5 이상 4.5 미만	0.4090	0.4104	0.1656	49	36	0	0.7347
4.5 이상	0.4879	0.4879	0.2465	26	20	0	0.7692

<표 9>를 보면 가중치 비가 작아짐에 따라 편향비율이 감소하는 것을 알 수 있다. 또한 가중치 비가 작을수록 $\bar{A}$ 는 감소하는 경향이 있다. 가중치 비에 따라 A_i 들이 통계적으로 유의한 차이가 있는지 살펴보기 위해 분산분석을 실시한 결과, F 값이 18.26이고 p 값이 0.0001보다 작아 유의한 차이가 있음을 알 수 있었다.



〈그림 11〉 가중치 비에 따른 A_i 의 분포



〈그림 12〉 가중치 비에 따른 $|A_i|$ 의 분포

그룹별 차이를 보기 위해, A_i 를 가지고 Fisher의 LSD 다중비교를 실시했으며, 그 결과는 <표 10>에 제시되어 있다. 가중치 비가 2.5 미만, 2.5 이상 3.5 미만, 3.5 이상으로 가중치 비가 증가할수록 A_i 가 유의하게 커진다는 것을 알 수 있으며, 이는 가중치 비가 감소하면 새누리당에 대한 치우침이 감소하여 보다 정확한 결과를 낸다는 것을 의미한다. $\bar{A}$ 가 대부분 양수이기 때문에 오차를 의미하는

$|A|$ 를 기준으로 한 분석에서도 비슷한 결론을 내릴 수 있었으며, 자세한 분석결과는 생략하였다.

<표 10> 가중치 비에 따른 A_i 의 LSD 다중비교 결과

가중치 비교	차이	p값	유의여부
1.5 미만-1.5 이상 2.5 미만	0.0197	0.7733	
1.5 미만-2.5 이상 3.5 미만	-0.0535	0.4438	
1.5 미만-3.5 이상 4.5 미만	-0.2096	0.0030	**
1.5 미만-4.5 이상	-0.2885	0.0001	***
1.5 이상 2.5 미만-2.5 이상 3.5 미만	-0.0732	0.0424	*
1.5 이상 2.5 미만-3.5 이상 4.5 미만	-0.2292	0.0000	***
1.5 이상 2.5 미만-4.5 이상	-0.3082	0.0000	***
2.5 이상 3.5 미만-3.5 이상 4.5 미만	-0.1561	0.0001	***
2.5 이상 3.5 미만-4.5 이상	-0.2350	0.0000	***
3.5 이상 4.5 미만-4.5 이상	-0.0789	0.0924	

한편, 상관분석을 해보면 가중치 비와 A_i 의 상관계수는 0.5168이고, p 값은 0.0001보다 작아 강한 양의 상관관계임을 알 수 있으며, 가중치 비와 $|A_i|$ 의 상관계수 역시 0.5150이고, p 값은 0.0001보다 작아 동일한 선형관계를 보인다. 따라서 상관분석을 통해 가중치 비가 커지면, 다시 말해 할당을 제대로 채우지 못 해 가중치 변동이 커지는 경우 여론조사의 정확성은 떨어진다는 것을 확인할 수 있다.

7. DB 활용조사를 제외한 응답률 및 가중치에 따른 예측오차 비교

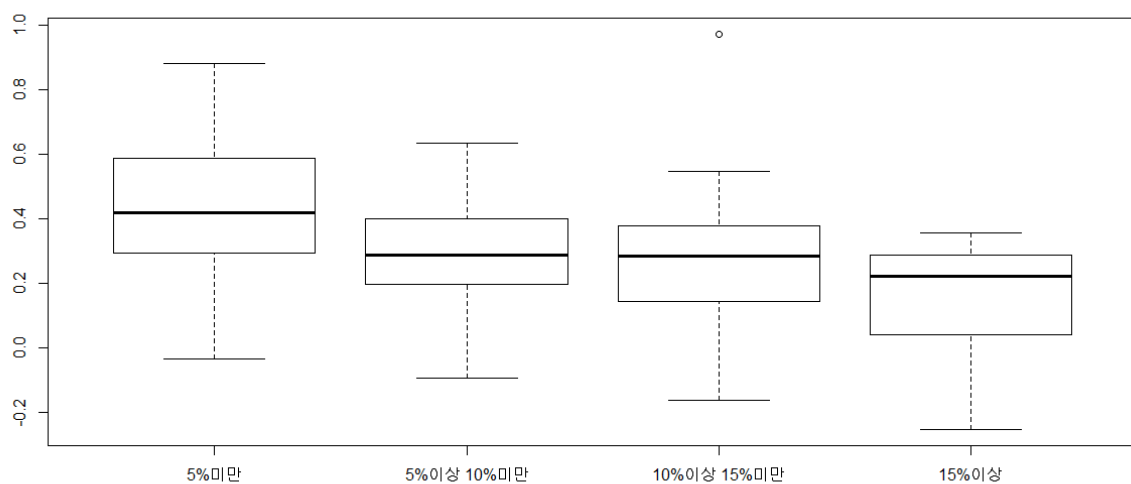
분석대상 여론조사 중에는 사전에 조사기관이 자체적으로 구축한 DB나 교통카드를 이용하는 사람들로 구축된 티머니 DB, 자발적 참여자를 중심으로 구축된 두잇서베이 DB 또는 마켓링크 DB 등을 이용한 조사들이 포함되어 있다. 이들

DB를 사용한 여론조사의 경우 응답률이나 가중치 비를 일반 전화조사에서의 응답률이나 가중치 비와 비교하는 것은 적절하기 않기 때문에 여기서는 사전에 구축해 놓은 이런 DB를 활용한 조사를 제외하고 응답률과 가중치 변동에 따른 오차를 분석해 보고자 한다.

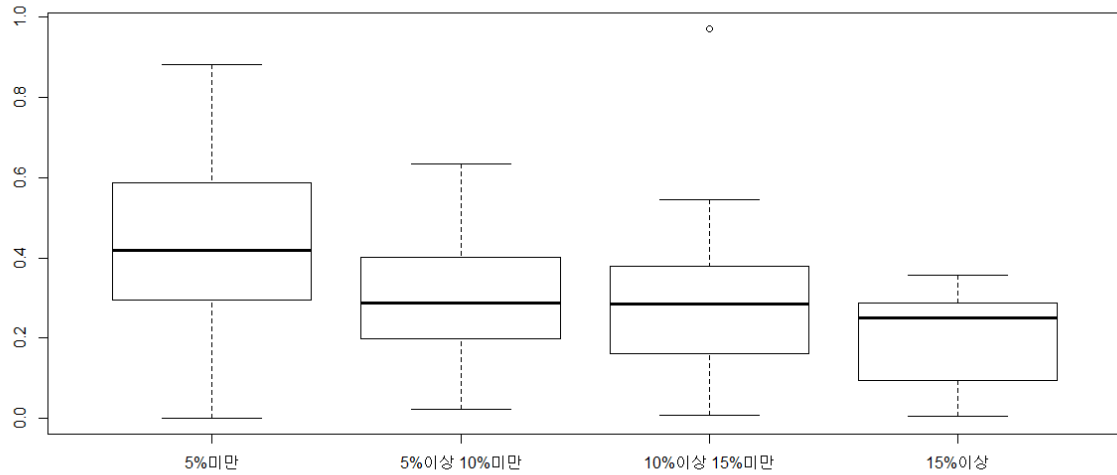
본 연구에서 사용한 분석대상 여론조사 중 DB를 활용한 여론조사는 총 74개였으며, 이를 제외한 나머지 129개의 여론조사를 대상으로 응답률 및 가중치에 따른 예측오차를 분석했다. DB를 사용한 여론조사를 제외한 후 응답률에 따른 편향과 예측오차는 <표 11>에 제시되어 있고, A_i 의 상자그림이 <그림 13>과 <그림 14>에 제시되어 있다.

<표 11> DB 활용조사를 제외한 응답률에 따른 편향 및 예측오차

응답률	$\bar{A}$	$ \bar{A} $	$s.d.(A)$	n	n^+	n^-	편향 비율
5% 미만	0.4385	0.4396	0.2138	60	46	0	0.7667
5% 이상 10% 미만	0.2779	0.2887	0.1788	41	18	0	0.4390
10% 이상 15% 미만	0.2685	0.2966	0.2663	17	6	0	0.3529
15% 이상	0.1454	0.2013	0.1881	11	3	0	0.2727



<그림 13> DB 활용조사를 제외한 응답률에 따른 A_i 의 분포



<그림 14> DB 활용조사를 제외한 응답률에 따른 $|A_i|$ 의 분포

DB 활용조사를 제외한 후에도 응답률이 높을수록 $\bar{A}$ 나 $|\bar{A}|$ 가 감소하는 것을 확인할 수 있다. A_i 에 대한 분산분석 결과, F 값이 9.483이고 p 값이 0.0001보다 작아 유의한 차이가 있다는 것을 알 수 있었으며, Fisher의 LSD 다중비교를 실시한 결과는 <표 12>에 제시되어 있다. $|A_i|$ 에 대한 분석에서도 거의 동일한 결과를 얻을 수 있었다. 또한 DB 활용조사를 제외한 후 응답률과 A_i 의 상관계수는 -0.4487, 응답률과 $|A_i|$ 의 상관계수는 -0.4185이고, p 값이 모두 0.0001보다 작아 강한 음의 상관관계를 가진다는 것을 알 수 있다. 결과적으로 응답률이 낮아지면 새누리당에 대한 치우침이 심하게 나타났고, 오차도 커진다고 말할 수 있다.

<표 12> DB 활용조사를 제외한 응답률에 따른 A_i 의 LSD 다중비교 결과

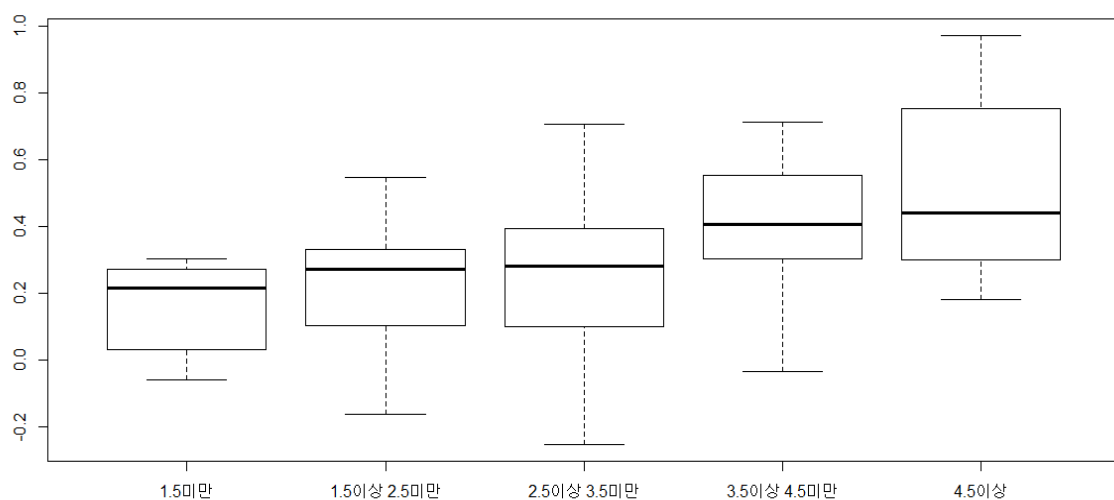
comparison	difference	p -value	sig.
5% 미만-5% 이상 10% 미만	0.1606	0.0002	***
5% 미만-10% 이상 15% 미만	0.1700	0.0037	**
5% 미만-15% 이상	0.2931	0.0000	***
5% 이상 10% 미만-10% 이상 15% 미만	0.0095	0.8757	
5% 이상 10% 미만-15% 이상	0.1325	0.0643	
10% 이상 15% 미만-15% 이상	0.1231	0.1307	

DB 활용조사를 제외한 후 가중치 비에 따른 편향과 예측오차는 <표 13>, A_i 와 $|A_i|$ 의 상자그림은 <그림 15>와 <그림 16>에 제시되어 있다.

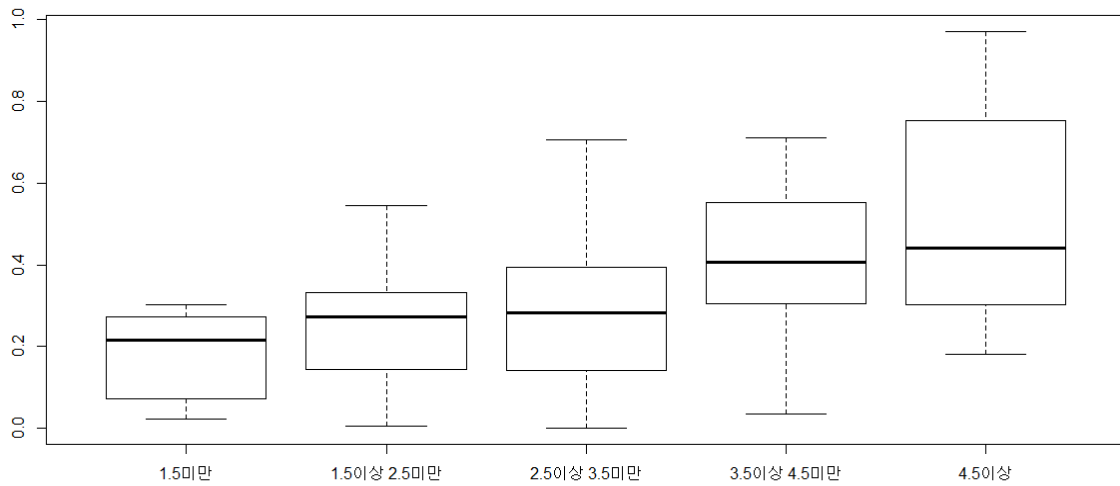
<표 13> DB 활용조사를 제외한 가중치 비에 따른 편향 및 예측오차

가중치 비	$\bar{A}$	$ \bar{A} $	$s.d.(A)$	n	n^+	n^-	편향 비율
1.5 미만	0.1612	0.1812	0.1423	8	0	0	0.0000
1.5 이상 2.5 미만	0.2223	0.2491	0.1839	29	10	0	0.3448
2.5 이상 3.5 미만	0.2651	0.2872	0.2255	27	13	0	0.4815
3.5 이상 4.5 미만	0.4129	0.4146	0.1721	41	31	0	0.7561
4.5 이상	0.5020	0.5020	0.2517	24	19	0	0.7917

<표 13>을 보면 편향비율이 가중치 비가 커짐에 따라 여전히 증가하는 경향이 있는 것을 확인할 수 있다. 또한 DB 활용조사를 제외한 후에도 가중치 비가 작을수록 $\bar{A}$ 나 $|\bar{A}|$ 가 감소하는 것을 확인할 수 있다.



<그림 15> DB 활용조사를 제외한 가중치 비에 따른 A_i 의 분포



〈그림 16〉 DB 활용조사를 제외한 가중치 비에 따른 $|A_i|$ 의 분포

〈표 14〉 DB 활용조사를 제외한 가중치 비에 따른 A_i 의 LSD 다중비교 결과

comparison	difference	p-value	sig.
1.5 미만-1.5 이상 2.5 미만	-0.0611	0.4498	
1.5 미만-2.5 이상 3.5 미만	-0.1039	0.2033	
1.5 미만-3.5 이상 4.5 미만	-0.2517	0.0016	**
1.5 미만-4.5 이상	-0.3408	0.0001	***
1.5 이상 2.5 미만-2.5 이상 3.5 미만	-0.0428	0.4294	
1.5 이상 2.5 미만-3.5 이상 4.5 미만	-0.1906	0.0002	***
1.5 이상 2.5 미만-4.5 이상	-0.2797	0.0000	***
2.5 이상 3.5 미만-3.5 이상 4.5 미만	-0.1478	0.0037	**
2.5 이상 3.5 미만-4.5 이상	-0.2369	0.0001	***
3.5 이상 4.5 미만-4.5 이상	-0.0891	0.0884	↯

A_i 에 대한 분산분석 결과, F 값이 10.17이고 p 값이 0.0001보다 작아 유의한 차이가 있다는 것을 알 수 있었으며, Fisher의 LSD 다중비교를 실시한 결과는 <표 14>에 제시되어 있다. $|A_i|$ 에 대한 분석에서도 거의 동일한 결과를 얻을

수 있었다. DB 활용조사를 제외한 후 가중치 비와 A_i 의 상관계수는 0.5154, 가중치 비와 $|A_i|$ 의 상관계수는 0.5089이고, p값이 두 경우 모두 0.0001보다 작아 강한 양의 상관관계를 가진다는 것을 알 수 있다. 결과적으로 DB 활용조사를 제외한 후에도 할당을 제대로 채우지 못 해서 가중치 변동이 커지게 되면 여론조사의 정확성이 떨어지게 된다는 결론을 내릴 수 있다.

IV. 전국단위 정당지지율 여론조사 분석

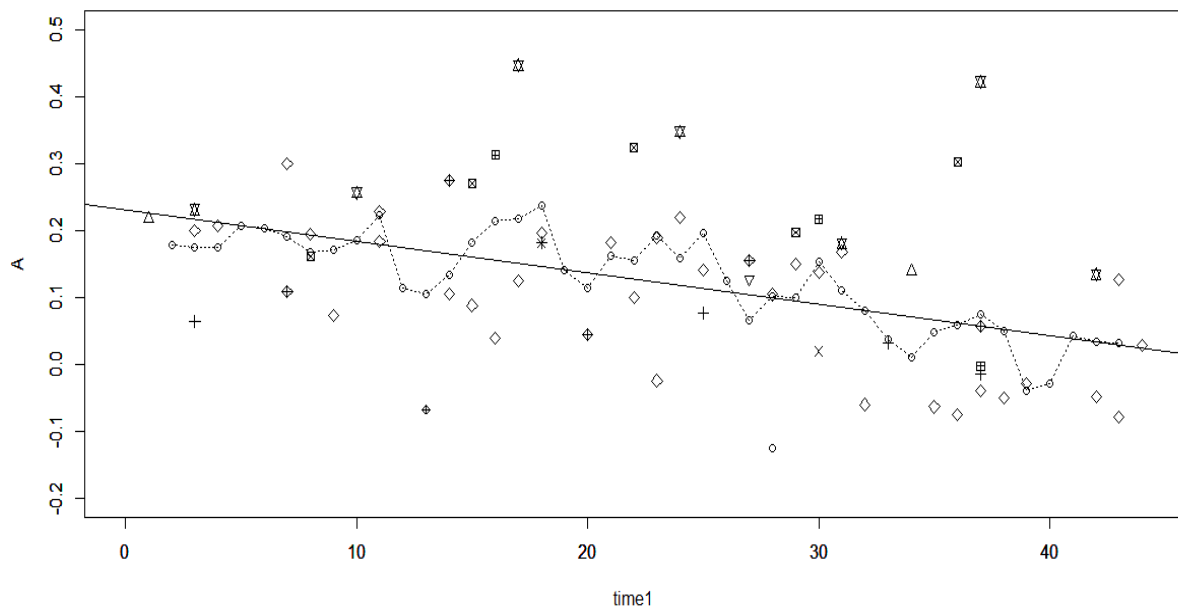
20대 국회의원선거 여론조사 중에는 정당별 지지율을 파악하기 위한 전국단위 조사들도 포함되어 있다. 실제 이번 선거의 경우 신당 창당 등을 포함해 정치 지형의 변화와 정당별 지지율에 영향을 줄 수 있는 여러 가지 사건들이 발생하였다는 점을 고려하여 투표결과와 비교했을 때, 정확성 관점에서 시간의 흐름에 따른 전국단위 정당지지율 여론조사 추이를 살펴보고자 한다. 특히 유권자들이 선거일에 임박해 지지 정당을 결정할 수 있고, 선거기간 막판에 지지 정당을 바꾸는 상황이 벌어질 수 있다는 점 등을 고려하면 일반적으로 선거일이 가까워질수록 여론조사의 정확성은 높아진다는 주장이 타당한 것인지 확인해 본다는 측면에서도 의미가 있을 것이다.

여론조사에서 예측한 정당별 지지율을 예측지지율로 하고, 20대 국회의원선거 비례대표 정당지지율을 실제 정당지지율로 보고, 척도 A 를 기준으로 여론조사의 정확성이 시간의 흐름에 따라 어떤 양상을 보이는지 분석하였다. 여기서는 3월 1일부터 선거 당일인 4월 13일까지 실시된 전국단위 정당지지율 여론조사 총 64개를 대상으로 하였다. 한 조사기관에서 일간, 주간, 주중으로 여론조사를 실시한 경우에는 일간 여론조사만을 포함하였다.

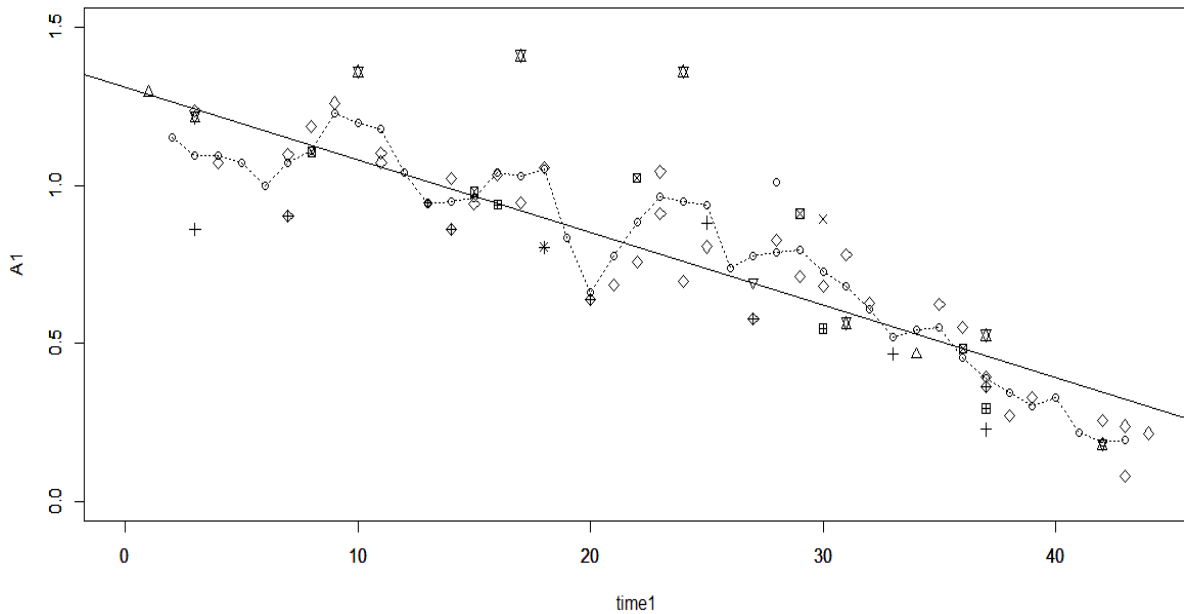
먼저, 새누리당과 더민주당의 지지율을 가지고 시간에 따른 여론조사 예측정확도 분석을 실시하였다. 시간의 흐름에 따른 변동을 보기 위해 3월 1일을 시작으로 일수를 계산하여 1에서 44의 값을 가지는 시간변수 t 를 생성하고, 시간 t 에

따른 척도 A 를 기준으로 한 새누리당과 더민주당에 대한 분석결과는 <그림 17>에 제시되어 있다.

각 점은 하나의 여론조사를 의미하며, 같은 모양을 가진 점은 같은 조사기관에서 실시한 여론조사를 나타낸다. 동그라미는 해당 일자 하루 전, 후의 날을 포함한 3일 동안 시행된 모든 여론조사에 대한 이동평균이고, 점선은 이동평균을 연결한 평활곡선이다. 실선은 예측정확도 A 를 반응변수, 시간변수 t 를 설명변수로 한 단순회귀직선이다. 적합된 직선의 식은 $\hat{A}_t = 0.2309 - 0.0047t$ 이며, 시간 t 의 회귀계수 -0.0047 의 p 값이 0.0001 보다 작아 매우 유의함을 알 수 있다. 적합된 단순회귀직선이 시간 t 가 커짐에 따라 0에 가까워지는 것으로 보아, 선거일이 다가올수록 여론조사의 정확도가 높아지 것을 알 수 있다. 하지만 이런 시간 흐름에 따른 변동을 오차의 감소로 보기보다는 시점별로 실제 정당별 지지율에 변동이 있었고 여론조사에 이런 변동이 반영되었던 것으로 해석할 수도 있다는 점에 유의할 필요가 있다.



<그림 17> 시간에 따른 척도 A 의 변화 추이(새누리당 & 더민주당)

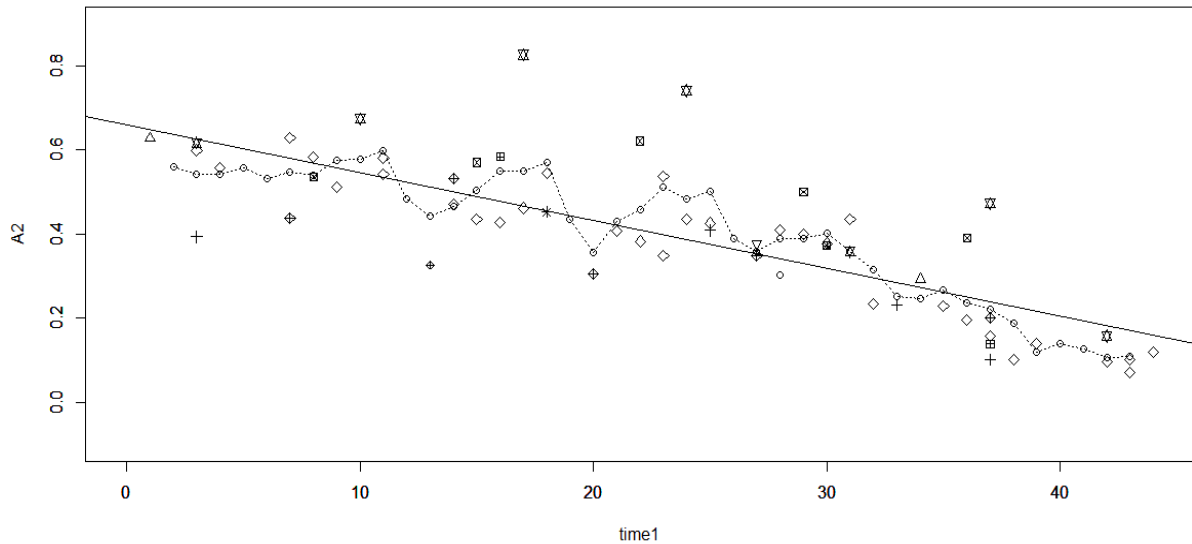


〈그림 18〉 시간에 따른 척도 A 의 변화 추이(새누리당 & 국민의당)

〈그림 18〉은 새누리당과 국민의당의 지지율을 가지고 시간에 따른 여론조사 예측정확도 추이를 분석한 결과이다. 시간 t 에 따른 예측정확도 분석 결과, 시간에 따른 예측정확도 추정식은 $\hat{A}_t = 1.3116 - 0.0230t$ 이며, 시간 t 의 회귀계수 -0.0230 의 p 값이 0.0001 보다 작아 매우 유의함을 알 수 있다. 적합된 단순회귀직선이 시간 t 가 커짐에 따라 0에 가까워지는 것으로 보아, 선거일이 다가올수록 여론조사 결과가 최종 투표결과에 가까워짐을 볼 수 있다. 앞서 새누리당과 더민주당의 지지율을 가지고 계산한 직선보다 기울기의 경사가 심한 것으로 보아 시간이 흐름에 따라 새누리당 지지가 더민주당보다는 국민의당으로 이전하는 경향이 더 강했다는 것으로 해석할 수 있을 것이다.

한편, 새누리당을 여당, 더민주당과 국민의당을 합쳐 야당으로 보고 여당과 야당의 지지율을 가지고 시간에 따른 여론조사 예측정확도 분석을 실시한 결과는 〈그림 19〉와 같다. 정당지지율을 여당과 야당으로 구분한 시간 t 에 따른 예측정확도 분석 결과, 시간에 따른 예측정확도 추정식은 $\hat{A}_t = 0.6594 - 0.0114t$ 이며, 시간 t 의 회귀계수 -0.0114 의 p 값이 0.0001 보다 작아 매우 유의함을 알 수 있다.

여기서도 선거일이 다가올수록 최종 투표결과를 기준으로 할 때 여론조사의 오차가 줄어든다는 것을 알 수 있으며, 이런 현상을 시간이 흐름에 따라 여당 지지율이 하락하고 야당 지지율이 상승했다는 것으로 해석할 수도 있을 것이다.



〈그림 19〉 시간에 따른 척도 A 의 변화 추이(여당 & 야당)

V. 결론

본 연구에서는 2016년 20대 국회의원선거 전화여론조사의 예측오차와 편향에 영향을 주는 요인에 대해 수도권 선거구에서 수행된 여론조사를 중심으로 살펴보았다. 또한 전국단위 정당별 지지율에 대한 여론조사의 정확성이 시간의 흐름에 따라 어떤 추이를 보였는지 정리하였다. 여심위 홈페이지에 등록된 자료를 바탕으로 후보자 등록일과 여론조사 공표일을 고려하여 3월 26일부터 4월 6일까지 수도권 선거구를 대상으로 실시된 여론조사를 분석대상으로 하였다.

Martin et al.(2005)이 제시한 정확성 척도 A 를 기준으로 분석해 본 결과, 경기, 인천, 서울의 모든 수도권 지역에서 새누리당 후보자를 과대추정하는 편향이

발생하였다는 것을 알 수 있었다. 조사기관별로 비교해 본 결과, 모든 조사기관에서 새누리당 후보자를 과대추정하는 편향이 발생하였으며 조사기관에 따라 예측오차 등에 있어 큰 차이를 보였다. 이는 박종희(2013)가 지적한 것처럼 동일한 조사방식을 사용하더라도 개별 조사기관의 고유한 특성에 의해 야기되는 조사기관 편향이 존재할 수 있다는 것을 의미한다.

또한 조사방법별로는 전화면접 방식이 ARS나 ARS와 전화면접을 비슷한 비율로 혼합한 방식에 비해 상대적으로 편향이나 오차가 적게 발생하는 것을 확인할 수 있었다. 조사시점별로는 4월3일 이전의 여론조사가 4월3일 이후의 여론조사보다 새누리당 후보자를 과대추정하는 편향이 심했던 것으로 나타났으며, 응답률이 증가하거나 가중치 변동이 작아질수록 편향과 예측오차가 감소하였다.

마지막으로 전국단위 정당지지율 여론조사를 대상으로 시간에 따른 정확도 분석을 실시한 결과, 시간의 흐름에 따라 여론조사 예측정확도 A 가 0으로 수렴해 선거일이 다가올수록 여론조사가 최종 투표결과와 일치하는 경향을 보인다는 것을 확인할 수 있었다. 이는 예측오차가 적어지는 것으로 볼 수도 있지만, 실제 시간이 흐름에 따라 유권자들의 정당지지율이 변화하고 여론조사에서 이런 변화를 일정 부분 설명해 주고 있는 것으로 해석할 수도 있을 것이다.

본 연구결과를 종합적으로 살펴보면, 김영원·황다솜(2014)의 연구에서 얻은 결과와 마찬가지로 비용이 좀 더 들더라도 ARS보다는 전화면접 방식을 통해 정확성을 높일 수 있고, 전통적으로 알려진 것과 같이 응답률을 높이고 할당을 제대로 채워 가중치 변동을 줄이는 것이 여론조사의 정확성을 확보하는 데 도움이 된다는 사실을 다시 한 번 확인할 수 있었다.

참고문헌

- 김영원·최윤정. 2011. “출구조사의 체계적인 예측 편향에 대한 분석: 2010년 지방선거 출구조사 중심으로.” 《조사연구》 12(3): 25-48.
- 김영원·황다솜. 2014. “2014년 지방선거 여론조사 전화조사방법에 따른 예측오차 및 편향.” 《조사연구》 15(4): 1-32.

- 박종희. 2013. “제18대 대선 여론조사에서 나타난 조사기관 편향.” 《조사연구》 14(1): 1-30.
- 최필선·민인식. 2013. “18대 대통령 선거 여론조사의 기관별 정확성 측정 및 비교.” 《조사연구》 14(2): 1-18.
- 윤지혜. 2016. “20대 총선 전화여론조사 사후분석.” 숙명여자대학교 대학원 석사 학위논문.
- Bolger, G., D.W. Davis, C. Franklin, R.M. Groves, P.J. Lavrakas, M.S. Mellman, P. Meyer, K. Olson, J.A. Selzer, M.W. Traugott, and C. Wlezien. 2009. *An Evaluation of the Methodology of the 2008 Pre-Election Primary Polls*. American Association for Public Opinion Research.
- Martin, E.A., M.W. Traugott, and C. Kennedy. 2005. “A Review and Proposal for a New Measure of Poll Accuracy.” *Public Opinion Quarterly* 69 (3): 342-369.
- Mokrzycki, M., S. Keeper, and C. Kennedy. 2009. “Cell-phone-only Voters in the 2008 Exit Poll and Implications for Future Noncoverage Bias.” *Public Opinion Quarterly* 73(5): 845-865.
- Sturgis, P., N. Baker, M. Callegaro, S. Fisher, J. Green, W. Jennings, J. Kuha, B. Lauderdale, and P. Smith. 2016. *Report of the Inquiry into the 2015 British General Election Opinion Polls*. London: British Polling Council and the Market Research Society.
- Traugott, M.W. and W. Christopher. 2009. “The Dynamics of Poll Performance During the 2008 Presidential Nomination Contest.” *Public Opinion Quarterly* 73(5): 866-894.

<접수 2016/08/18, 게재확정 2016/08/27>