

## KBS 세대 인식조사

### 주관적 계층의식과 세대 및 성별 간 관계

메모(2021년 6월 30일)

김석호(서울대학교 사회학과)  
임동균(서울대학교 사회학과)  
하상응(서강대학교 정치외교학과)

이 메모는 KBS 세대 인식 조사 자료를 분석하여 얻은 “주관적 계층 의식과 세대 및 성별 간 관계”에 대한 설명을 담고 있습니다. 지난 2021년 6월 20일 KBS 시사기획 <창>에 조사 내용의 일부가 보도된 이후, 보다 상세한 결과가 KBS 홈페이지에 네 차례에 걸쳐 실렸습니다. 이 중에서 응답자의 주관적 계층 의식, 세대, 성별, 그리고 이타적인 태도 간 관계를 묘사하는 그래프에 대한 관심이 SNS와 인터넷 게시판 등에서 매우 뜨거웠습니다. 그래프가 보여주고자 한 바는 **50대 남녀 및 20-34세 여성과 달리, 20-34세 남성은 자신이 소속한 계층이 높다고 생각할 수록 “우리 사회의 어려운 사람들을 위해 내가 가진 것을 나누어 주고 싶다”는 생각을 덜 한다는 것이었습니다.** 이를 검증하기 위해 활용한 통계기법은 변수들 간 교차표를 만들어 작성한 기술통계가 아니라 대학교/대학원 중급통계 시간에 다루는 순서형 로짓과 이항 로짓이라는 점을 먼저 밝힙니다. 적지 않은 수의 SNS 사용자들이 이 그래프에 담긴 의미가 정확하게 무엇인지, 이 그래프는 어떻게 그려졌는지, 이 그래프가 학술적인 발견을 제시하는 것 말고 다른 정치적인 의미를 두고자 조작된 것은 아닌지에 대한 의구심을 제기하였습니다. 이 메모는 이러한 의구심을 해소하기 위해 작성되었습니다.

(자료 분석 과정을 명료한 사회과학 용어로 설명하기 위해 어쩔 수 없이 사회조사 방법론 및 사회통계학에서 통용되는 학술 용어를 사용합니다. 따라서 이 분야 사전 지식이 없으신 일반인들은 이해하시기 어려울 수도 있습니다. 이 점 미리 양해 구합니다.)

#### 1. 자료와 변수들

분석에 사용한 설문 자료는 KBS에서 주관하고 여론조사회사 <한국리서치>에서 수행한 온라인 설문 자료입니다. 청년세대(20-34세)와 소위 “586”이라고 불리는 50대 간 상호 인식이 어떠한지를 다양한 측면에서 보기 위한 설문이었습니다. 그래서 일반적인 여론조사와는 달리 20-34세 응답자 600명, 50대 응답자 600명, 총 1,200명을 접촉해 설문을 수행하는 방식을 택했습니다. 응답자 접촉 방식 및 표본 구성 방식은 <한국리서치>에서 수행하는 여론조사의 관행을 그대로 따랐습니다.

그래프에 담긴 변수들은 다음과 같습니다.

종속변수: 남에 대한 이타적인 태도 (도움 의향)

다음 문항에서 얻은 정보를 종속변수로 사용했습니다. 4 점 척도 문항입니다.

**Q47. 다음의 질문에 대한 귀하의 생각을 말씀해주세요. [랜덤][척도][STEPBYSTEP]**

	매우 그렇다	그렇다	그렇지 않다	전혀 그렇지 않다
	1	2	3	4
1. 기회만 있다면 우리 사회의 어려운 사람들을 위해 내가 가진 것을 나누어 주고 싶다				

독립변수: 주관적 계층 인식

다음 문항에서 얻은 정보를 독립변수로 사용했습니다.

**Q1. 한국사회의 최하층을 1로 하고, 최상층을 10으로 한다면 귀하는 어디에 속한다고 생각하십니까? [단일척도]**

최하층	<---	---	---	---	---	---	---	---	---	--->	최상층
1	2	3	4	5	6	7	8	9	10		

조절변수: 세대/성별 (모두 네 집단 - 20-34 세 남성; 20-34 세 여성; 50 대 남성; 50 대 여성)

## 2. 관심 변수들 간의 관계(기술통계량)

독립변수인 주관적 계층 인식과 세대 간 관계는 <표 1>과 같습니다. 자신을 최상층(9 혹은 10)에 위치시킨 응답자의 수는 0 에 가깝습니다. 이것은 별로 놀라운 결과가 아닙니다.

**<표 1> 세대/성별 과 주관적 계층 인식 간 관계(응답자수)**

	2034 남성	2034 여성	50 대 남성	50 대 여성	Total
① 최하층	6	3	9	4	22
②	8	10	14	8	40
③	45	55	56	42	198
④	65	69	61	60	255
⑤	74	82	69	94	319
⑥	68	38	52	51	209
⑦	34	22	32	24	112
⑧	13	8	10	10	41
⑨	0	0	1	3	4
⑩ 최상층	0	0	0	0	0
Total	313	287	304	296	1,200

독립변수인 주관적 계층 인식과 종속변수 간의 관계는 <표 2>와 같습니다.

<표 2> 주관적 계층인식과 도움 의향 간 관계(응답자 수)

	(1) 매우 그렇다	(2) 그렇다	(3) 그렇지 않다	(4) 전혀 그렇지 않다	Total
① 최하층	4	10	3	5	22
②	4	26	9	1	40
③	15	131	47	5	198
④	22	178	52	3	255
⑤	23	215	74	7	319
⑥	16	140	46	7	209
⑦	11	80	18	3	112
⑧	4	24	12	1	41
⑨	0	3	0	1	4
⑩ 최상층	0	0	0	0	0
Total	99	807	261	33	1,200

마지막으로 세대/성별과 종속변수 간의 관계는 <표 3>과 같습니다.

<표 3> 세대/성별과 도움 의향 간 관계(응답자 수, 괄호 안 퍼센티지)

	(1) 매우 그렇다	(2) 그렇다	(3) 그렇지 않다	(4) 전혀 그렇지 않다	Total
20-34 남성	35 (11.18)	173 (55.27)	90 (28.75)	15 (4.79)	313 (100.00)
20-34 여성	29 (10.10)	183 (63.76)	65 (22.65)	10 (3.48)	287 (100.00)
50 대 남성	22 (7.24)	226 (74.34)	52 (17.11)	4 (1.35)	304 (100.00)
50 대 여성	13 (4.39)	225 (76.01)	54 (18.24)	4 (1.35)	296 (100.00)
Total	99 (8.25)	807 (67.25)	261 (21.75)	33 (2.75)	1,200 (100.00)

<표 3>을 보면 전체적으로 20-34 세 응답자들이 50 대 응답자에 비해 도움 의향에 대한 부정적인 답변을 많이 한 것을 알 수 있습니다. 20-34 남성의 경우 그 비율이 28.75%(그렇지 않다) 더하기 4.79%(전혀 그렇지 않다), 20-34 여성의 경우 22.65%(그렇지 않다) 더하기 3.48%(전혀 그렇지 않다)이기 때문에, 20-34 남성의 경우 돕지

않겠다는 의향을 밝힌 비율이 약 34% 가량 되고, 20-34 여성의 경우에는 그 비율이 약 26% 정도 됩니다. 한편 50 대 응답자의 경우, 그 비율이 남녀 불문하고 20%가 채 되지 않습니다.

### 3. 분석 1: 순서형 로짓

종속변수가 4 점 척도이고 응답 범주에 순서가 있기 때문에 순서형 로짓(ordered logit)을 적용하는 것이 적절하다고 판단해 위에 제시된 변수들만 넣고 분석했습니다. 독립변수인 주관적 계층 인식이 종속변수에 주는 영향이 세대/성별 집단에 따라 다른지를 보기 위해 세대/성별 집단과 독립변수 간 교차항도 모형에 포함시켰습니다. 결과는 다음과 같습니다.

여기서 해석의 편의를 위해 종속변수의 변수값을 재코딩했습니다. 원래 "(1) 매우 그렇다; (2) 그렇다; (3) 그렇지 않다; (4) 전혀 그렇지 않다"로 되어있던 변수값을 "(1) 전혀 그렇지 않다; (2) 그렇지 않다; (3) 그렇다; (4) 매우 그렇다"로 바꾸었습니다.

```
. ologit revQ47_1 ib(0) .gen4##c.sss_m, nolog
```

```
Ordered logistic regression          Number of obs   =       1,200
                                   LR chi2(7)         =       24.90
                                   Prob > chi2        =       0.0008
Log likelihood = -1071.4906         Pseudo R2       =       0.0115
```

revQ47_1	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
gen4						
2034여 성	.2354811	.1753572	1.34	0.179	-.1082127	.5791749
50대 남 성	.4642374	.171574	2.71	0.007	.1279586	.8005162
50대 여 성	.30196	.1705872	1.77	0.077	-.0323848	.6363049
sss_m	-.2470625	.080378	-3.07	0.002	-.4046005	-.0895244
gen4#c.sss_m						
2034여 성	.3267762	.1208139	2.70	0.007	.0899854	.5635671
50대 남 성	.394108	.1101871	3.58	0.000	.1781452	.6100708
50대 여 성	.3468803	.1156548	3.00	0.003	.120201	.5735595
/cut1	-3.377097	.2024822			-3.773955	-2.980239
/cut2	-.9068801	.123326			-1.148595	-.6651657
/cut3	2.685597	.1528342			2.386047	2.985147

위의 Stata 결과는 '20-34 남성' 집단을 기준 범주로 놓은 결과를 보여줍니다. 독립변수인 주관적 계층 인식(sss)이 높으면 높을수록 남을 돕겠다는 정도가 줄어들고 있음을 음(-)의 값을 갖는 회귀계수를 통해 알 수 있습니다.<sup>1</sup> 이 결과는 유의수준 0.01 하에서 통계적으로 유의미한 결과입니다. 그런데 주관적 계층 인식과

<sup>1</sup> 주관적 계층 인식변수('sss')는 평균중심화 하였습니다. 'gen4' 더미 변수들 해석시 유의 바랍니다. 평균중심화 여부는 결과에 실질적 영향을 미치지 않았습니다.

세대/성별 간 상호작용항의 회귀계수를 보면('gen4#c.sss\_m') 모두 양(+)<sup>2</sup>의 값을 갖고, 유의수준 0.01 하에서 통계적으로 유의미한 결과입니다. 이것은 20-34 남성 집단과 비교해 볼 때, 다른 세대/성별 집단들에게서는 주관적 계층 인식이 높아질 수록 남을 돕겠다는 정도가 덜 줄어들음을 의미합니다. (즉, 기준 범주인 20-34 남성에게서는 주관적 계층 인식의 회귀계수가 [-]인데, 다른 세대/성별 집단에게는 그 [-]을 상쇄하는 [+]값이 작동한다는 이야기입니다. 이 [+]값이 바로 다른 세대/성별 집단과 주관적 계층 인식 간 상호작용항에 딸린 회귀계수의 부호입니다.) 이를 통해 20-34 남성들의 경우 주관적 계층 인식이 높아질수록 남을 돕겠다고 답변할 확률이 줄어들 것이란 예상을 할 수 있습니다.

이 결과는 통제변수들을 고려해 보아도 크게 바뀌지 않습니다. 교육수준(고졸 이하, 대재 이상), 거주지역, 혼인 상태(미혼, 배우자 있음, 사별, 이혼), 월가구 소득(0~300 만원 미만, 300 만원~600 만원 미만, 600 만 원 이상), 부채 제외 가구 순자산(5000 만원 미만, 5000 만원~3 억 원 미만, 3 억 원) 등을 고려한 결과가 아래에 있습니다.<sup>2</sup>

```
. ologit revQ47_1 ib(0).gen4##c.sss_m i.marital i.educ_d i.region i.inc3 i.wealth, nolog
```

```
Ordered logistic regression          Number of obs      =       1,200
                                     LR chi2(31)         =       48.13
                                     Prob > chi2         =       0.0256
Log likelihood = -1059.8757          Pseudo R2          =       0.0222
```

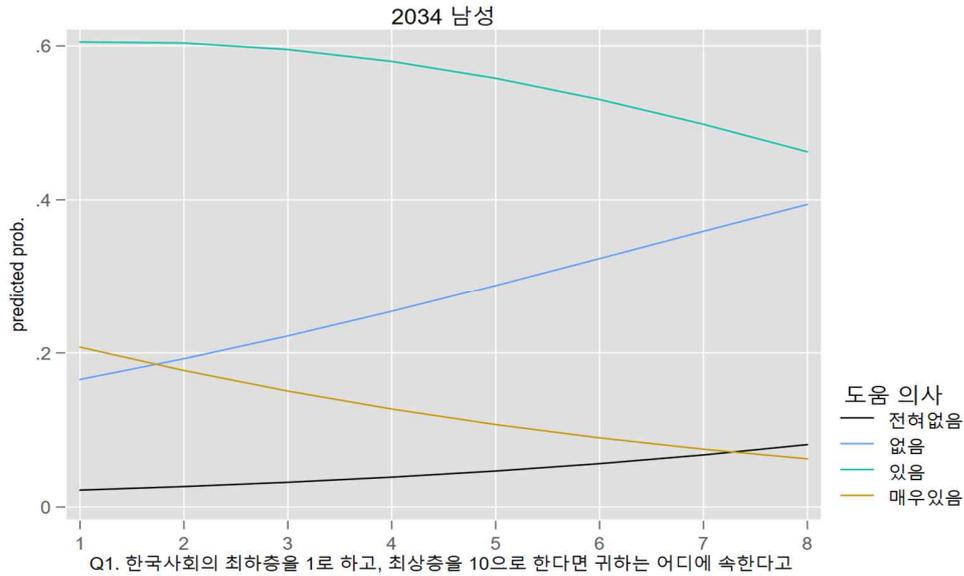
revQ47_1	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
gen4					
2034여 성	.2260613	.1805779	1.25	0.211	-.1278648 .5799874
50대 남 성	.5134672	.2327365	2.21	0.027	.0573122 .9696223
50대 여 성	.2885793	.2433126	1.19	0.236	-.1883046 .7654633
sss_m	-.2541502	.0832859	-3.05	0.002	-.4173876 -.0909127
gen4#c.sss_m					
2034여 성	.3485979	.1230261	2.83	0.005	.1074711 .5897246
50대 남 성	.4521571	.1133638	3.99	0.000	.2299682 .674346
50대 여 성	.4154633	.1184994	3.51	0.000	.1832087 .6477179

(통제변수가 보이는 하단 생략 - 유의한 변수 없음)

그런데 위에서 본 바와 같이 순서형 로짓 결과는 해석하기가 매우 까다롭습니다. 그래서 이 결과를 **예측확률(predicted probability)**을 계산하여 그래프로 시각화 하면 예를 들어 아래와 같은 방식으로 표현할 수 있습니다. <그림 1>~<그림 4>에서 확인하실 수 있습니다.

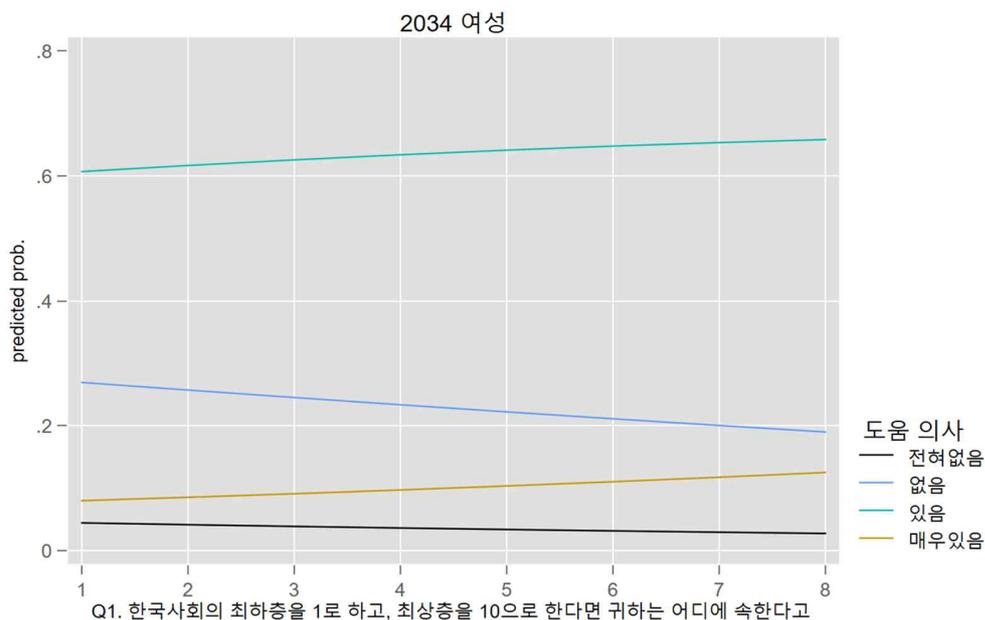
<sup>2</sup> 여기서 sss 의 효과를 보는 데 있어 과연 이와 같은 통제변수들을 투입하는 것이 필요한지 여부에 대해서도 기존 연구들을 바탕으로 토론할 수 있겠지만, 이 리서치 메모는 KBS 와 회의를 위해 만들었던 그래프를 만드는데 사용했던 모델에 대해 설명하는 것이 목적이기 때문에 그 때 사용했던 모델을 그대로 보여드리는 바입니다.

<그림 1> 주관적 계층 인식이 도움 의향에 미치는 효과 (예측확률: 20-34 남성의 경우)

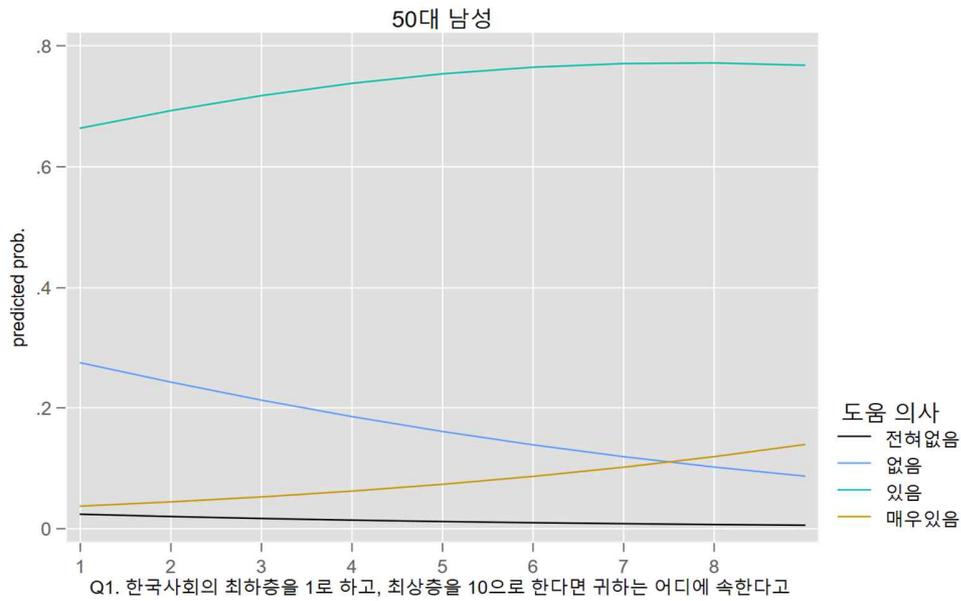


위에서 시행한 순서형 로짓 결과, 20-34 남성의 경우 주관적 계층 인식이 높아짐에 따라 남을 도울 의사가 “전혀 없음”이거나 “없음”일 확률이 커짐을 확인할 수 있습니다. 예를 들어 독립변수 값이 2 인 20-34 남성의 경우 (다른 조건이 일정하다면) 남을 도울 의사가 “없음”을 선택할 확률이 20%(세로축에 표시한 예측확률 값 = 0.2)임에 비해 독립변수 값이 8 인 20-34 남성은 남을 도울 의사가 “없음”을 선택할 확률이 40%(세로축에 표시한 예측확률 값 = 0.4)임을 확인할 수 있습니다.

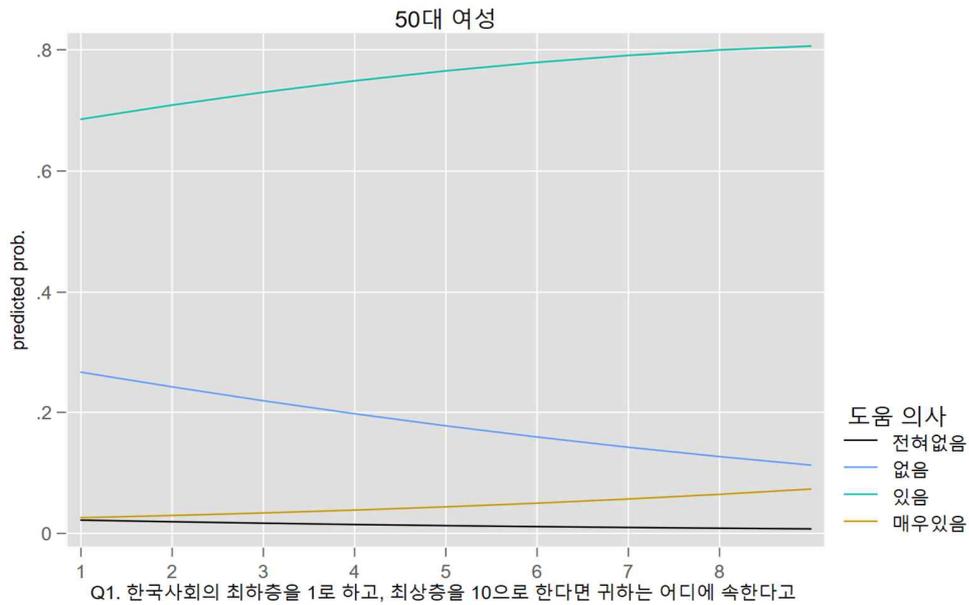
<그림 2> 주관적 계층 인식이 도움 의향에 미치는 효과 (예측확률: 20-34 여성의 경우)



<그림 3> 주관적 계층 인식이 도움 의향에 미치는 효과 (예측확률: 50 대 남성의 경우)



<그림 4> 주관적 계층 인식이 도움 의향에 미치는 효과 (예측확률: 50 대 여성의 경우)



그런데 20-34 남성과 달리 다른 세대/성별 집단에서는 주관적 계층 인식이 높을 수록 남을 돕겠다고 답할 확률이 높아짐을 확인할 수 있습니다. 위의 네 개의 그래프는 순서형 로짓 결과를 활용해 만든 것입니다.

위의 결과를 조금 더 친절하게 보여주기 위해 세대/성별 집단 별 순서형 로짓을 각각 돌린 후, 주관적 계층 인식이 도움 의사에 미치는 영향이 어떠한지를 정리하면 <표 4>와 같습니다. 표에서 나타나듯이, 20-34 남성의 경우에만 주관적 계층 인식이 도움 의향에 미치는 효과(회귀계수)가 음수(-)임을 볼 수 있습니다.

<표 4> 세대/성별 별 주관적 계층 인식이 도움 의향에 미치는 효과(회귀계수 및 표준오차)

	통제변수 미포함	통제변수 포함
20-34 남성	-.195** (.075)	-.190* (.085)
20-34 여성	.071 (.086)	.100 (.107)
50 대 남성	.178* (.082)	.181 (.113)
50 대 여성	.136 (.094)	.314** (.120)

#### 4. 분석 2: 이항 로짓

하지만 <표 4>에 제시된 정보 역시 언론보도를 통해 직관적인 해석이 가능한 정보가 아니기 때문에, 종속변수의 정보량은 감소하겠지만 직관적인 해석을 위해, 종속변수의 응답 범주 중 도움을 줄 의사가 “있다”는 범주들과 “없다”는 응답범주들을 각각 합쳐 이항 변수(0 or 1)로 만들었습니다. 결과적으로 도움을 줄 의사가 있다(1)가 전체의 75.5%, 없다(0)가 전체의 24.5%인 변수('help')가 구성됩니다. 새롭게 만든 종속변수를 이용해 이항 로짓 분석을 하면 결과가 다음과 같습니다.

```
. logit help ib(0).gen4##c.sss_m, nolog
```

```
Logistic regression                Number of obs    =      1,200
                                   LR chi2(7)        =      38.33
                                   Prob > chi2         =      0.0000
Log likelihood = -648.9635         Pseudo R2       =      0.0287
```

help	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
gen4						
2034여 성	.3211817	.1833136	1.75	0.080	-.0381063	.6804697
50대 남 성	.8315754	.1989873	4.18	0.000	.4415674	1.221583
50대 여 성	.6851225	.191859	3.57	0.000	.3090858	1.061159
sss_m	-.1906103	.08074	-2.36	0.018	-.3488577	-.0323629
gen4#c.sss_m						
2034여 성	.2350591	.1246786	1.89	0.059	-.0093066	.4794247
50대 남 성	.4301141	.1239398	3.47	0.001	.1871966	.6730316
50대 여 성	.3265062	.1293477	2.52	0.012	.0729893	.5800231
_cons	.7260482	.1231373	5.90	0.000	.4847035	.9673928

앞의 순서형 로짓 분석에서 포함한 통제변수들을 넣고 이항 로짓 분석을 재차 수행하면 그 결과가 다음과 같습니다.

```
. logit help ib(0).gen4##c.sss_m i.marital i.educ_d i.region i.inc3 i.wealth, nolog
```

```
Logistic regression                Number of obs    =      1,200
                                   LR chi2(31)       =      54.75
                                   Prob > chi2         =      0.0053
Log likelihood = -640.75465         Pseudo R2       =      0.0410
```

help	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
gen4						
2034여 성	.3367746	.1899187	1.77	0.076	-.0354592	.7090083
50대 남 성	.9153467	.2645324	3.46	0.001	.3968727	1.433821
50대 여 성	.7420944	.2722212	2.73	0.006	.2085508	1.275638
sss_m	-.1956711	.0848887	-2.31	0.021	-.36205	-.0292922
gen4#c.sss_m						
2034여 성	.2279556	.1277711	1.78	0.074	-.0224711	.4783824
50대 남 성	.4736766	.1288325	3.68	0.000	.2211695	.7261837
50대 여 성	.3725113	.1326492	2.81	0.005	.1125235	.632499

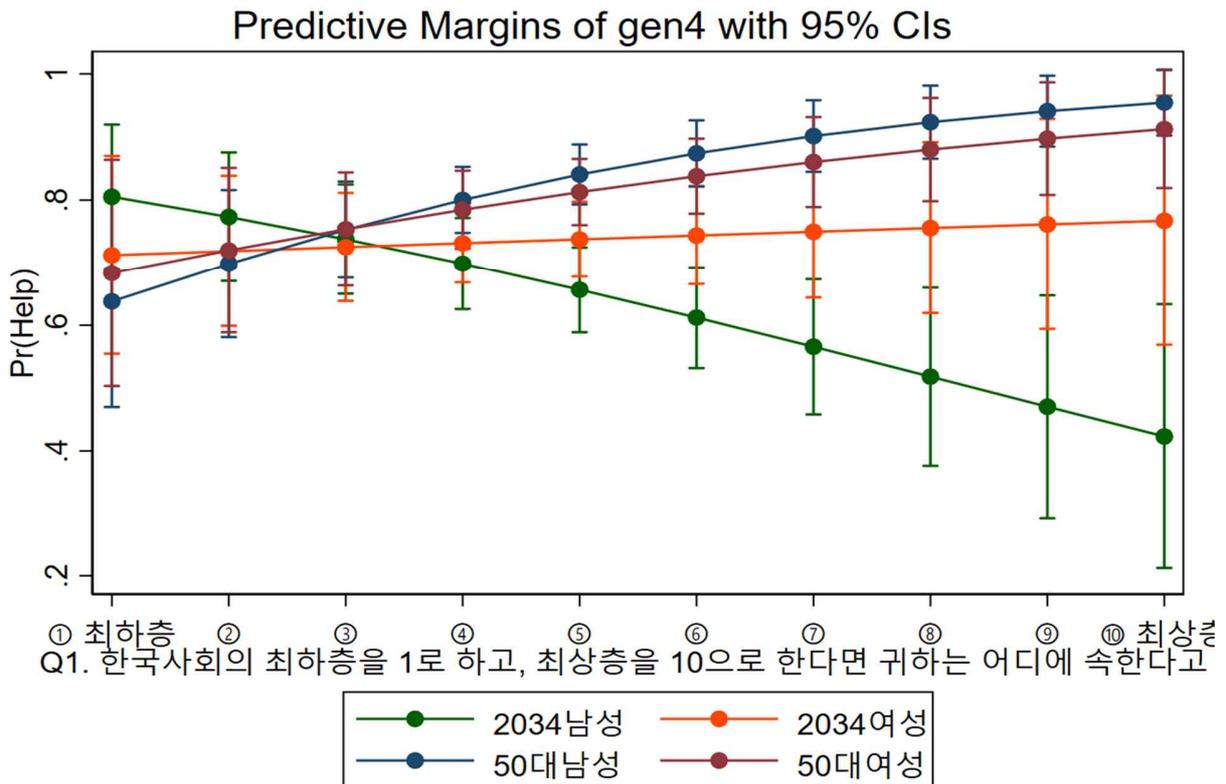
(통제변수가 보이는 하단 생략 - 유의한 변수 없음)

이항 로짓 분석 결과는 앞의 순서형 로짓 분석 결과와 유사합니다. 다만, 순서형 로짓 결과에서는 20-34 남성과 20-34 여성 간의 차이도 유의수준 0.01 하에서도 통계적으로 유의미하게 나온 반면, 로짓 분석에서는 20-34 남성과 20-34 여성 간 차이가 유의수준 0.10 하에서만 통계적으로 유의미합니다. 통제변수 없이 이항 로짓 분석을 한 결과 이들 두 집단 간의 차이의 p 값은 0.059, 통제변수 포함하고 이항 로짓 분석을 한 결과 이들 두 집단 간의 차이의 p 값은 0.074 이기 때문입니다. 일반적으로 양측 검정을 수행할 때 통계적 유의미성 기준으로 유의수준 0.05 를 사용하기 때문에 이러한 p 값은 “애매한 유의미성(borderline significance)”라고 부를 수 있습니다.

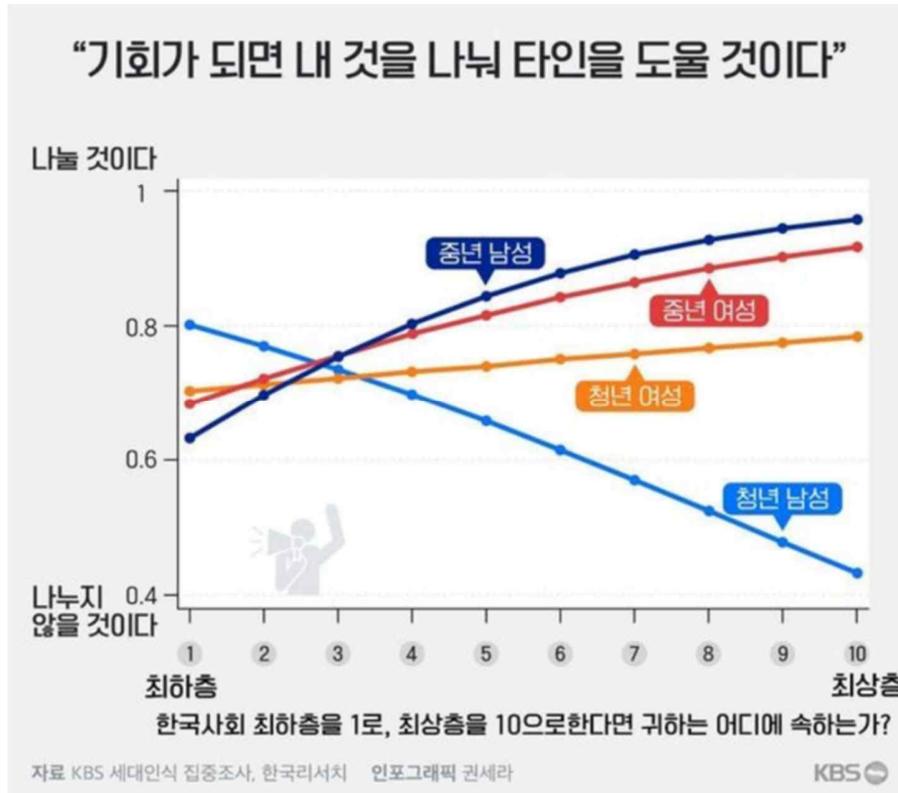
(하지만 중요한 점은 종속변수를 원래 그대로의 모양인 4 점 척도로 순서형 로짓 분석을 했을 때에는 20-34 남성 집단이 다른 세 개의 세대/성별 집단과 통계적으로 유의미한 차이를 보였다는 것입니다.)

이 로짓 결과를 그래프로 구현하면 <그림 5>와 같습니다. 통제변수를 고려한 모형에 기반한 그래프이고 **이것이 바로 KBS 가 홈페이지에 올린 그래프입니다.** 단 한가지 다른 점은 여기 보이는 그래프에서는 95% 신뢰구간을 표시했다는 것입니다. 위의 로짓 결과에서 20-34 남성과 20-34 여성 간 차이의 p 값이 0.074 였기 때문에 이들 두 집단 간의 95% 수준에서의 신뢰구간은 약간 겹칩니다(신뢰구간이 겹치는 것이 곧 통계적 차이가 없음을 의미하지는 않습니다. 겹치는 정도에 따라 결정됩니다.). **하지만 20-34 남성과 50 대 남성 및 50 대 여성 간 차이는 x 값이 5 일때부터 확연히 드러납니다.**

<그림 5> 주관적 계층 인식이 도움 의향에 미치는 효과: 예측확률 (이항 로짓 분석 결과)



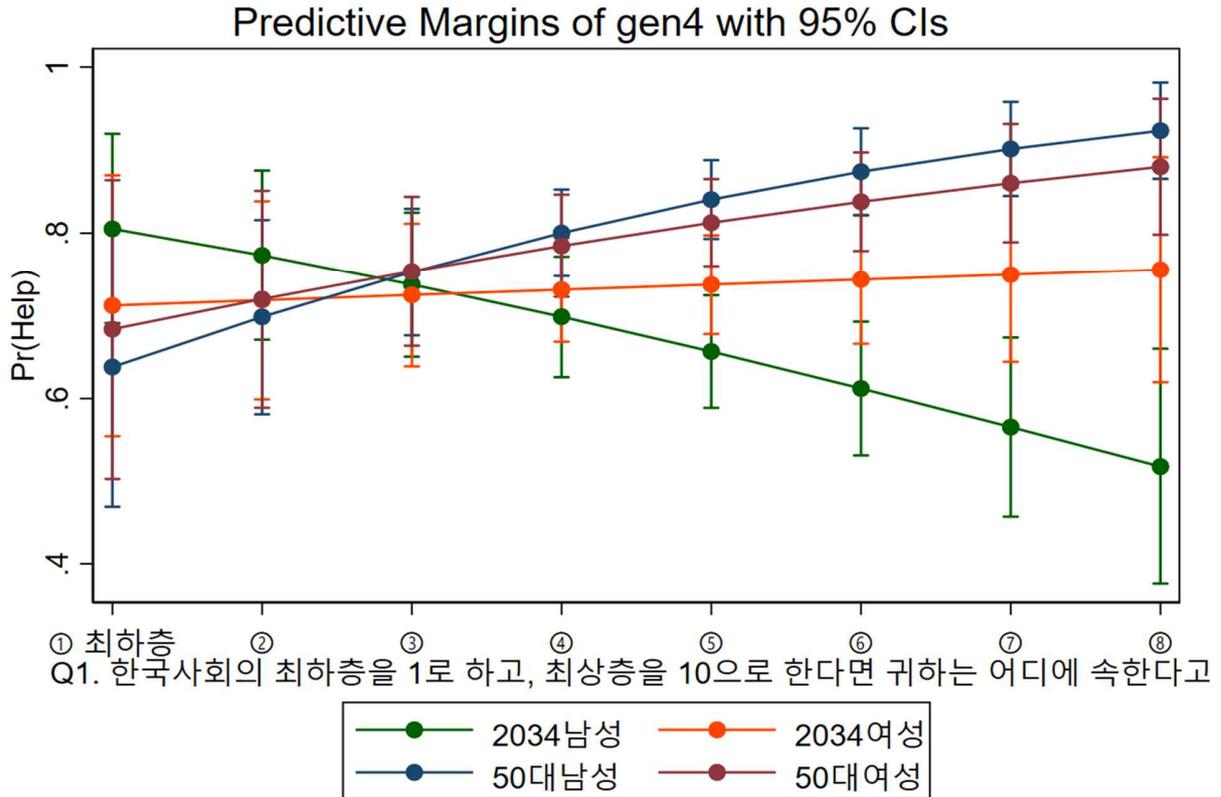
<그림 5>와 KBS 에서 올린 그래프를 비교해 보십시오. (위 그림의 경우 신뢰구간을 표기하지 않을 시 y 축은 Stata 에서 0.4~1 로 자동 조정되어 산출됩니다.)



여기서 강조해야 할 점 두 가지는 첫째, 이 조사의 원래 초점과 목표는 매우 분명하게 (586 에 초점을 맞춘) ‘세대 인식’을 탐색하는 것이었고, 세대 간 차이와 인식을 파악하는 점이었던 것입니다. 2034 남성의 상대적 특징 또한 그러한 맥락에서 주목한 것입니다. 둘째, 연구자들이 가진 문제의식은 상업적인 세대론에서 주장하는 것과는 달리 인위적이고 자의적인 세대 구분에 의한 세대 간 차이를 부각시키는 것보다는 세대 집단 내 다양성과 이질성에 주목해야 각 집단이 가진 문제가 무엇인지 구체적으로 드러낼 수 있다는 것이었습니다. 그런 차원에서 세대별 주관적 계층 인식에 따른 도움 의향 등을 포함한 여러 인식과 행위의 차이를 살펴보게 된 것입니다.

주관적 계층의식 값이 10 인 응답자가 없는 것과 9 인 응답자가 매우 소수인 점을 고려해서 x 축의 범위를 1-8 로 조정한 후 그린 그래프는 다음과 같습니다.

<그림 6> 주관적 계층 인식이 도움 의향에 미치는 효과: 예측확률 (이항 로짓 분석 결과 - x 축 재조정)



## 5. 정리

KBS 세대 인식조사 보도 관련하여 SNS 에서 논란이 된 그래프는 (1) 연구진이 이미 확인한 순서형 로짓 분석 결과에 기반하여, (2) 가독성 높은 그래프를 만들기 위해 추정 모형을 이항 로짓으로 바꾸어 분석을 수행하는 과정에서 생산된 것입니다. (이 또한 사실상 '최종' 그래프라기 보다는 분석결과를 KBS 에 소개하고 회의를 하기 위한 자료로 만들었던 것입니다.) 자료가 담은 정보를 보다 정확하게 반영하는 결과는 순서형 로짓 분석을 통해 확인할 수 있습니다. 보도된 이항 로짓 분석 결과는 순서형 로짓 분석 결과와 매우 유사하긴 하나, 종속변수의 정보가 약간 다르기 때문에 자료의 정보를 정확히 반영해 주지 않습니다. 결국 연구진이 분석작업을 하며 여러 그래프들을 만들 때 언론 보도를 염두에 두고 가독성이 높은 그래프를 생산하기 위해 불필요한 별도의 작업을 한 것이 분석 결과를 정확하게 전달하지 못하게 된 큰 이유라고 하겠습니다. 이 메모가 문제의 그래프를 둘러싼 불필요한 논쟁을 종식시키는 데에 도움이 되었으면 합니다.

연구진들이 확실하게 이야기 할 수 있는 내용은 다음과 같습니다.

이 설문 자료의 **대표성이 담보된다**는 전제하에, 회귀 모형에 포함된 통제변수의 값을 **평균값에 고정시켰을 때**, 주관적 계층 인식이 높아질 수록 20-34 세 남성은 **평균적으로** 다른 집단(20-34 세 여성, 50 대 남성, 50 대 여성)에 비해 통계적으로 유의미하게 남을 돕고자 하는 의향이 감소한다. 이 결과는 순서형 로짓 결과에서는 분명하게 나오지만, 종속변수를 재코딩하여 이항 로짓을 사용했을 때에는 20-34 세 남성과 20-34 세 여성 간 차이가 유의수준 0.10 에서만 통계적으로 유의미하게 나온다. (이항 로짓을 사용했을 때 20-34 세 남성과 50 대 남성 혹은 여성 간 차이는 유의수준 0.05 에서 통계적으로 유의미하다.)

(회귀분석 결과를 해석하는 과정에는 여러가지 전제조건이 달립니다. 위의 문단에서 빨간색으로 표시된 부분이 결과 보고 시 보통 생략되는 전제조건들입니다. 이와 같은 전제조건들이 생략되는 이유는 회귀분석에 능한 연구자들이 서로 그 전제조건을 암묵적으로 공유하고 있기 때문입니다.)

하지만 설문조사를 분석하여 결과를 내고 해석하는 사회과학 경험 연구가 갖는 한계로부터 이 결과가 완전히 자유롭지는 않습니다. 몇 가지 주의점들과 제안하고 싶은 내용을 정리해 보겠습니다.

### 1. 표본의 대표성 문제

이 분석에서 사용된 자료가 한국의 성인남녀를 정확하게 대표하는 표본인지는 알 길이 없습니다. 대표성을 담보하기 위해 여러 기법이 동원되지만, 실질적으로 모집단을 정확하게 대표하는 설문조사 표본은 존재하지 않는다고 보는 것이 맞습니다. 하지만 명명백백하게 대표성이 망가진 표본은 판단할 수 있습니다. 가령 이 분석에서 독립변수인 주관적 계층 인식의 1 값부터 10 값에 해당하는 사람 각각 100 명을 뽑아 총 1,000 명의 응답자로 구성된 표본을 만든다면, 이것은 명명백백하게 대표성이 없는 표본이 됩니다. 왜냐하면 상식적으로 보았을 때 모집단(한국의 전체 성인남녀) 수준에서, 자신을 최상층이라고 생각하거나 최하층이라고 생각하는 사람의 비율이 자신을 중산층이라고 생각하는 사람의 비율보다 작을 것이기 때문입니다.

### 2. 결과 해석의 문제

이 메모에 제시된 결과는 단 하나의 설문조사에서 도출된 결과입니다. 만약 같은 방식으로, 같은 설문 도구를 사용해, 별도의 1,200 명을 모은 제 2, 제 3 의 설문조사가 있다면, 거기서부터 나오는 결과는 이 조사의 결과와 다를 가능성이 농후합니다. 선거여론조사에서 특정 후보 지지도가 여론조사 기관마다 다른 이유와 유사합니다. 따라서 여기에 제시된 연구 결과는 “참고할 만한 증거(suggestive evidence)”에 불과합니다. 이 발견이 중요하다면, 후속 연구를 통해 재확인될 필요가 있습니다. 그리고 혹시라도 이 연구와 유사한 연구가 과거에 있었다면 그 연구 결과와의 비교도 필요합니다. 사회과학, 특히 심리학, 사회심리학, 정치심리학, 보건학 등에서 사용되는 메타분석(meta-analysis)이라는 아이디어가 이러한 문제의식에서 비롯된 것입니다.

### 3. “왜” 이런 결과가?

이 결과를 받아들인다고 해도, 왜 20-34 세 남성에게서만 독특한 태도가 발견되었는지에 대한 궁금증은 여전히 남습니다. 안타깝게도 이 질문에 대한 답을 얻기 위해서는 별도의, 본격적인 연구가 필요합니다. 즉, 이 질문에 대한 답을 찾기 위해서는 독립변수와 종속변수 간 관계를

엮어주는 매개변수(mediating variable)를 찾고, 독립변수 → 매개변수 → 종속변수를 잇는 인과고리를 규명해야 하는데 이 작업이 생각보다 만만치 않습니다. 이 과정에서 (기사에 소개되었듯) 연구진이 기자에게 말한 “논문거리”라는 표현이 등장했습니다. 보도된 그래프의 차이가 왜 이렇게 나왔으며 어떻게 설명할 수 있는가의 질문을 받고 연구진의 답변은 분석하기 전에는 “모른다” 였습니다. (그러므로 논문거리라는 의미.) 사실 지금도 우리는 그 이유를 여러 가지로 추측할 뿐, 가설을 세워 검증해보기 전에는 “왜”라는 질문에 정확한 답변을 하기 어렵습니다. 따라서 이에 대한 어떤 의미부여(이대남에 대한 비난 등)도 해서는 안된다는게 연구진의 입장입니다. 이런 저런 추측과 우리의 직관에 따르면 많이 가진 사람이 다른 사람에게 베풀고자 하는 경향이 클 것입니다. 하지만 이러한 직관에 반하는 경험적 근거를 찾아온 연구도 있습니다.<sup>3</sup> 이 연구에 따르면 사람은 오히려 가질 수록 덜 베푸는 동시에 비도덕적인 행동을 할 가능성이 높다고 합니다. 이 연구 전통에서 확인한, 혹은 확인하고자 하는 매개 요인들을 살펴보는 후속 연구를 수행한다면 이 설문조사에서 얻은 결과의 해석이 보다 풍부해 질 수 있을 것 입니다. **아직은 왜 20-34 세 남성의 행태가 특이하게 나타났는지에 대한 속시원한 설명은 없습니다. 따라서 “20-34 세 남성들이 왜 이런 행태를 보이는가”라는 질문에 대한 설득력 있는 해답을 찾기 전에는 20-34 세 남성의 태도와 행동에 대한 규범적, 도덕적 판단은 할 수도 없고 당연히 자제해야 합니다.**

#### 4. 소위 “세대론”에 대하여

최근 한국 사회에서 활발하게 논의되고 있는 세대론에 대해서도 이 연구결과는 시사점을 제공해 줍니다. 우선 “세대 간 차이”에만 초점을 맞추지 말고 “세대 내 차이”에도 관심을 가져야 한다는 교훈을 줍니다. 이 연구에서 밝혀진 20-34 세 남성과 여성의 차이가 그 예입니다. 하지만 이 연구 말고도 20 대 혹은 청년세대에서 남녀 차이가 생긴 원인, 과정, 그 결과물에 대한 연구는 무수히 많이 존재합니다. 이 연구는 젠더 기준으로 나뉘는 세대 내 차이 뿐만 아니라, 상대적으로 덜 관심을 받아왔던 계급(혹은 계급 인식) 기준으로 나뉘는 세대 내 차이도 중요할 수 있다는 점을 시사해 줍니다.

#### 5. 사회과학 경험연구와 미디어

사회과학 경험 연구 결과가 미디어에서 어떻게 다루어져야 하는지의 문제에 대한 고민이 필요합니다. 이 메모가 작성되게 된 이유도 연구진들이 미디어 종사자들과 자료 분석 결과를 공유하는 과정에서 충분한 논의가 진행되지 않았기 때문입니다.

---

<sup>3</sup> Piff, P. K., Kraus, M. W., Côté, S., Cheng, B. H., & Keltner, D. (2010). Having less, giving more: the influence of social class on prosocial behavior. *Journal of personality and social psychology*, 99(5), 771-784; Piff, P. K. (2014). Wealth and the inflated self: Class, entitlement, and narcissism. *Personality and Social Psychology Bulletin*, 40(1), 34-43.

연구진은 KBS로부터 <시사기획 창>의 586 및 세대 인식과 관련된 방송을 위한 설문 문항 기획과 및 조사자료 분석과 자문을 의뢰 받고, 한국리서치의 설문조사를 통해 수집된 설문 자료에 대한 분석을 수행하였습니다. 200 여개 문항으로 이루어진 설문 자료에서 의미있는 패턴을 발견하기 위해 다량의 분석과 매우 많은 수의 그래프를 산출하였고, 이를 제작진과의 미팅 자료로 만들었습니다. 이번에 주목을 받은 그래프는 그 회의를 위해 제공된 자료들 중 하나였습니다.

다만 그 이후 기사로 보도되기 전까지 제작진과 자문단 사이에 어떤 그래프를 선정할 것인지, 선정된 그래프를 소개하기 전에 확인할 사항이 없는지 등과 관련된 별도의 논의가 이루어지지 못했던 아쉬움은 있습니다. (하지만 이것이 일반적인 관행에 비춰봤을 때 이례적인 케이스도 아니며, KBS 에서 연구 결과를 소개함에 있어 어떠한 '특별한' 의도를 가지고 있었다거나 연구 결과 소개를 왜곡하였다고 생각하지도 않습니다.)

자료분석으로부터 최종적으로 보도되는 과정에 있어서의 절차적 신중함, 연구자들이 분석 결과의 내용을 훼손하지 않으면서도 쉬운 용어와 쉬운 방식으로 일반 대중에게 다가갈 수 있는 방법이 무엇인지에 대한 고민을 해야 되겠습니다. 이 부분이 부족했음에 대해 연구진은 깊은 반성을 하고 있으며 이러한 소란을 유발한 부족함에 대해 사과 드립니다. 동시에 사회과학 경험 연구 결과를 받아 보도하는 미디어에서도 데이터 전문 인력을 육성하여 학계와의 소통을 원활히 할 장치를 마련했으면 합니다.

## 6. 기술적인 문제들

이 밖에도 (1) 주관적 계층 인식이라는 변수가 갖는 문제점, (2) 종속변수에서 발견될 수 있는 사회적 바람직성 편향의 가능성, (3) 이항 로짓 혹은 순서형 로짓 분석 시 모형 적합도의 계산과 해석, (4) 통제변수인 소득 및 자산과 독립변수 간 다중공선성 문제 등과 관련된 질문들이 있었습니다. 이 내용은 연구진이 분석 결과를 학술논문 혹은 연구노트의 형식으로 만드는 과정에서 다룰 계획입니다.

## 7. 자료의 공개 여부

연구진은 자료를 즉각적으로 투명하게 공개하기를 희망합니다. 하지만 연구진이 자료 소유권을 갖고 있지 않기 때문에 자료 소유권자인 KBS 와 자료 공개 시점과 방식을 논의하는 중입니다. 공개의 시기와 방식이 결정되자마자 오해의 소지를 없애기 위해 사회과학 자료를 전문적으로 아카이빙해 제공하는 기관에 의뢰해 공개하도록 하겠습니다.