Online Appendix for "To Denounce, or Not To Denounce: Survey Experiments on Diplomatic Quarrels"

Shoko Kohama, Kazunori Inamasu, and Atsushi Tago

Contents

Identification of "satisficing" respondents . . . 2
DK responses . . . 3
Supplemental analysis for Experiment 1 (Online full version) . . . 4
The full results of Experiment 2 . . . 10
Note on the frequencies of scrambling . . . 12
Balance tables . . . 13

References **22**

List of Figures

1 Experiment 1: "Don't know" and message heterogeneity. . . . 3
2 Experiment 1: OLS and OLS with Entropy Balancing. . . . 8
3 Experiment 1: OLS by Respondents Correctly/Incorrectly Identifying the Context. . . . 9

List of Tables

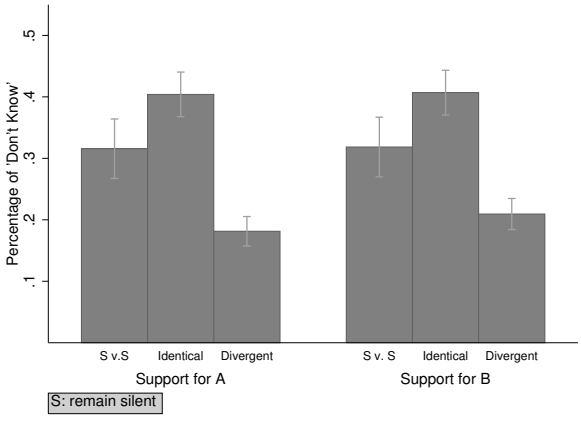
1 Experiment 1: OLS and Ologit Analysis . . . 7
2 Experiment 2: ROK . . . 10
3 Experiment 2: US . . . 11
4 Experiment 1: Balance table: full sample . . . 14
5 Experiment 1: Balance table: full sample, cont. . . . 15
6 Experiment 1: Balance table: sample without DK . . . 16
7 Experiment 1: Balance table: sample without DK, cont. . . . 17
8 Experiment 2, ROK: balance table . . . 18
9 Experiment 2, ROK: balance table, cont. . . . 19
10 Experiment 2, US: balance table . . . 20
11 Experiment 2, US: balance table, cont. . . . 21

Identification of "satisficing" respondents

In Experiment 1, to identify possible "inattentive" or "satisficing" respondents under imperfect control in an internet-based survey, we incorporated two "Screeners" in our experiment (Berinsky et al., 2014; Krosnick, 1991). First, we asked each respondent to "select the rightmost choice for this question" within a sequence of multiple-choice questions in the middle of the questionnaire, by which 156 respondents were entrapped. Secondly, we measured each respondent's answer time on each page of the experiment. 125 respondents spent less than 30 seconds on the page of the main questionnaire consisting of a description of a scenario and fifteen multiple-choice questions. "Screened" people overlap and hence the total number of participants who were possibly "inattentive" was 224. They appear to be younger than others, although it is unclear if they report their demography attentively. We confirmed that our findings remain unchanged even if we narrow our sample to "attentive" respondents. Considering the small number of "inattentive" respondents, it is hard to draw a reasonable inference on them. In Experiment 2, we were not allowed to include a trap question and then applied the 30 seconds criterion.

DK responses

It is known that people answer "I don't know" for various reasons. While DK responses can be induced truly by lacking of an underlying attitude, they may also include false negatives where people with an attitude refuse to give an opinion (Gilljam & Granberg, 1993). Moreover, they may be resulted from *satisficing* by respondents with low cognitive skills or lacking proper attention (Krosnick et al., 2002). Therefore, Beatty & Herrmann (1995) list four potential reasons for DK responses to attitudinal questions; namely, uncertainty, ambivalence, inaccessibility of information, and low motivation. DK responses in our experiment should be associated with uncertainty and ambivalence for the following rationale: we allowed respondents to choose "don't want to answer" besides "don't know"; our results remain unchanged even if we omit seemingly *satisficing* respondents (See the previous section); we presented a hypothetical scenario and our attitudinal questions presuppose no background knowledge other than those provided in the scenario. Thus, respondents choosing "don't know" in our experiment appear to be those who do not form a firm attitude or who feel undecided as to whether to support Country A (or B).



Note: Whiskers indicate 95% C.I.

Figure 1. Experiment 1: "Don't know" and message heterogeneity.

Supplemental analysis for Experiment 1 (Online full version)

This section addresses issues in our study and provides supplemental analyses. One of the biggest concerns arises from whether the sample in our study is unrepresentative. Although we were able to reach a diverse pool of the Japanese population, participants voluntarily preregistered to the survey company and hence may share particular characteristics.

To explore this problem, we reexamine the experimental outcome through regression analyses with and without modifying the bias caused by sampling. For correction, we employ the entropy balancing technique and reweight the survey sample to known features of the target population (Hainmueller & Xu, 2013). Specifically, we use age, gender, education level, income and residential location to capture the characteristics of the Japanese population.[1]

The upper panels of Figure 2 report the results of OLS models and the lower panels presents those using entropy balancing. The figure indicates the average effect of each treatment on levels of public support for Country A or B with a 95% confidence interval, setting *silence v. silence* as a baseline. Because the estimate for the baseline is roughly 1.6, that is somewhere between "oppose" and "somewhat oppose" for all models, the average treatment effect of 0.5 means that respondents on average "somewhat oppose" and 1.5 implies that the treatment alters their attitude to "somewhat support."[2]

The OLS analysis confirms our arguments addressed in the previous sections and they remain unchanged even if we take into account the bias in our sample. First of all, the upper-left panel of the figure demonstrates that both *self-promotion* and *denouncement* have significantly positive effects on public support for Country A, comparing with *silence*. Similarly, the upper-right panel shows that *self-promotion* and *denouncement* significantly increase public support, while *silence* doesn't. The OLS analysis with entropy balancing (the lower panels) generates the same results.

Secondly, *denouncement* is an even more powerful means of increasing popularity than *self-promotion*. Hypothesis tests (Wald tests) verify that the average treatment effect of *denouncement v. silence* is significantly higher than that of *self-promotion v. silence* ($p = 0.00$). Because Country B's strategy is constant in the two treatments, the positive difference is caused by Country A's *denouncement*. Similarly, the average treatment effect of *denouncement v. self-promotion* is significantly higher than that of *self-promotion v. self-promotion* ($p = 0.00$). The results remain unchanged in the case of OLS with entropy balancing.

[1]For age, gender, education level and income, we referred to The Comparative Study of Electoral Systems (2015). For residential location, we created a binary variable for each of five regions, including Hokkaido-Tohoku, Kanto-Koshinetsu, Chubu, Kinki, Chugoku-Shikoku and Kusyu-Okinawa. We used statistics from Statistics Bureau, Ministry of Internal Affairs and Communications (2014).

[2]Note that the predicted values derived from the OLS model can be biased due to the nonlinearity of the dependent variable.

Thirdly, the effect of *denouncement* diminishes in the case of mutual accusation. *denouncement* is particularly effective if Country B remains *silent* or *self-promote*. However, the effect of *denouncement v. denouncement* is significantly smaller than that of *denouncement v. silence* ($p = 0.00$) and *denouncement v. self-promotion* ($p = 0.00$). Entropy balancing does not alter our conclusion.

The fourth point is that *denouncement* undermines the public support for the attacked. According to the upper-right panel, both *self-promotion v. self-promotion* and *denouncement v. self-promotion* have significantly positive effects on the public support for Country B. Comparing the size of their effects, the effect of the former treatment is significantly larger than that of the latter ($p = 0.00$). When we adjust the sampling bias with entropy balancing (the lower-right panel), moreover, the effect of the latter becomes insignificant. This implies that Country A's verbal attack undermines the popularity of Country B, given that Country B adopts *self-promotion* in both cases.

The last but not least important finding is that *silence* protects countries from verbal attacks. The upper-right panel reports insignificant effects of *self-promotion v. silence* and *denouncement v. silence* in comparison with *silence v. silence*. This suggests that public support is not significantly affected by Country A's strategy as long as Country B remains *silent*. Again, the analysis with entropy balancing confirms this finding.

In addition to the sampling bias, we must also address that some respondents may be aware of the original context of our experimental scenario, i.e. the scrambling cases between China and Japan in 2014. In fact, about a half of the respondents were able to give the right answer to the post-experimental, multiple-choice question regarding the original context of the scenario. They tend to be old, male, more educated, and more conservative.[3] Some of them revealed that they answered the questions, having particular countries in mind.

Although the scenario mimicking the actual incidents lends ecological validity to our study, it poses some challenges at the same time. One concern is that respondents who were able to correctly identify the original context are presumably more knowledgeable and hence, more sensitive to the tenor of campaigns than others. This deserves serious attention, considering that only a part of population consumes information on low salience disputes between foreign countries. Another problem arises if respondents answer experimental questions, consciously or unconsciously assuming that either Country A or B is Japan. In such a case, their responses may not measure "foreign" citizens' opinion as we intended.

Therefore we conduct supplemental analyses, dividing our sample into two groups that consist of respondents giving the correct answer and those giving the incorrect answer, respectively (Figure 3). Overall, there is little difference between the two groups, except for the modest impact of messages and the particularly small effect of

[3]Although we do not report the results of the logit analysis here, it is included in our replication file.

denouncement v. denouncement among respondents giving the incorrect answer. We confirmed that all findings highlighted so far remain unchanged across the groups.

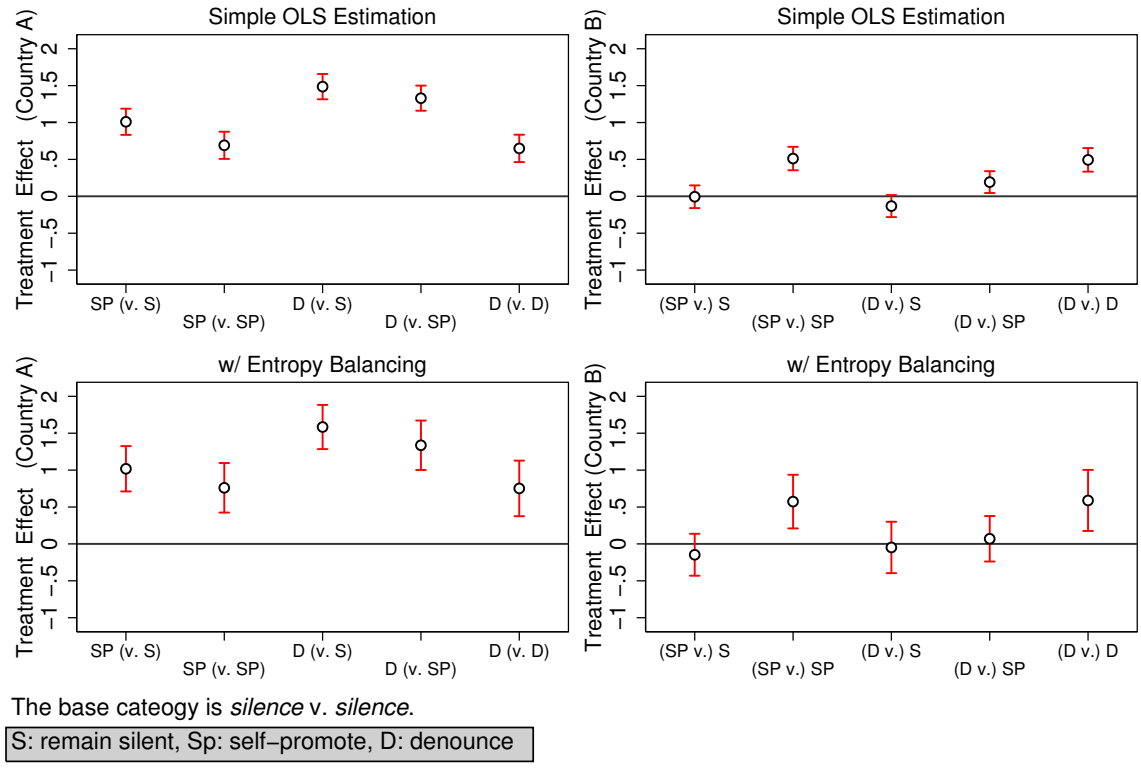
Highlighting a dissimilarity in the nearly identical results of the two groups, we find that those lacking political knowledge are less capable of perceiving the difference in nuances of *self-promotion* and *denouncement*. The effect of messages is moderate among respondents giving the incorrect answer, especially when it comes to *denouncement*. Besides, *denouncement v. denouncement* exhibits only a small effect on Country B's popularity. Its effect is not significantly larger than that of *denouncement v. self-promotion*, whereas the difference between the two treatments is significant in all other models and (sub)samples. Taken together, *self-promotion* and *denouncement* do not have distinct impacts on the attitude of respondents who failed to identify the original context. This is presumably because respondents lacking political knowledge do not recognize the different connotations behind *self-promotion* and *denouncement*.

Table 1
Experiment 1: OLS and Ologit Analysis

	OLS		OLS w. EB		Ologit w. EB	
	(1)	(2)	(3)	(4)	(5)	(6)
	Support A	Support B	Support A	Support B	Support A	Support B
main						
SP v. S	1.010***	-0.006	1.018***	-0.147	1.697***	-0.277
	(0.090)	(0.078)	(0.157)	(0.145)	(0.297)	(0.320)
SP v. SP	0.691***	0.512***	0.761***	0.574**	1.264***	1.068**
	(0.094)	(0.081)	(0.171)	(0.185)	(0.317)	(0.347)
D v. S	1.487***	-0.132	1.584***	-0.048	2.709***	-0.180
	(0.087)	(0.076)	(0.153)	(0.178)	(0.322)	(0.359)
D v. SP	1.330***	0.193*	1.337***	0.070	2.262***	0.200
	(0.087)	(0.075)	(0.171)	(0.157)	(0.335)	(0.317)
D v. D	0.649***	0.494***	0.752***	0.589**	1.237***	1.021*
	(0.094)	(0.081)	(0.192)	(0.211)	(0.371)	(0.424)
Constant	1.563***	1.602***	1.597***	1.693***		
	(0.064)	(0.055)	(0.100)	(0.119)		
cut1						
Constant					0.391	0.115
					(0.220)	(0.247)
cut2						
Constant					1.315***	1.526***
					(0.213)	(0.260)
cut3						
Constant					2.868***	2.695***
					(0.246)	(0.315)
R-squared	0.202***	0.074***	0.212***	0.082***		
N	1481	1450	1469	1439	1469	1439

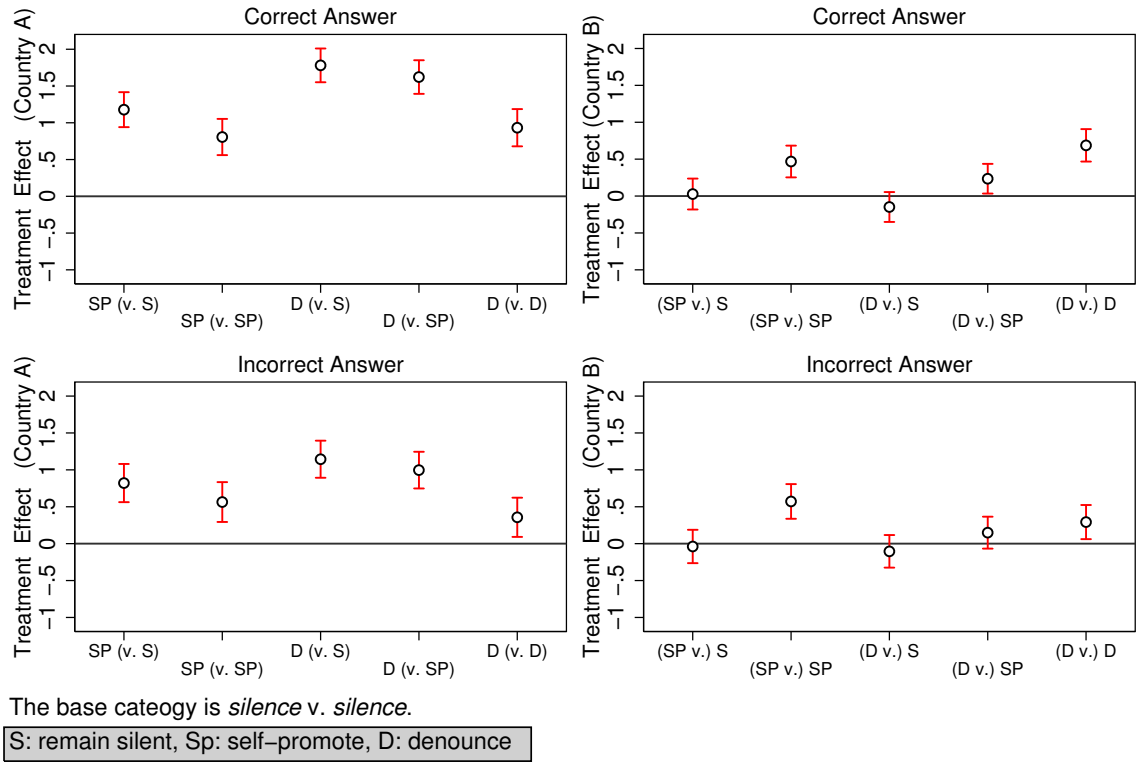
The base category is silence v. silence.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$



Note: Whiskers indicate 95% C.I.

Figure 2. Experiment 1: OLS and OLS with Entropy Balancing.



Note: Whiskers indicate 95% C.I.

Figure 3. Experiment 1: OLS by Respondents Correctly/Incorrectly Identifying the Context.

The full results of Experiment 2

Table 2
Experiment 2: ROK

	OLS		Attentives†		Ologit	
	(1)	(2)	(3)	(4)	(5)	(6)
	CHN	JPN	CHN	JPN	CHN	JPN
China v. Japan						
SP v. S	0.228*	0.086	0.263**	0.103	0.481*	0.429*
	(0.092)	(0.064)	(0.095)	(0.061)	(0.187)	(0.216)
S v. SP	0.064	-0.019	0.152	0.020	0.130	-0.040
	(0.092)	(0.063)	(0.097)	(0.061)	(0.191)	(0.224)
SP v. SP	0.028	0.012	0.123	0.093	0.074	0.168
	(0.093)	(0.065)	(0.098)	(0.063)	(0.191)	(0.224)
D v. S	0.353***	0.055	0.416***	0.123*	0.723***	0.368
	(0.092)	(0.063)	(0.094)	(0.061)	(0.189)	(0.215)
S v. D	-0.005	0.123*	0.042	0.157**	-0.008	0.512*
	(0.092)	(0.063)	(0.097)	(0.061)	(0.189)	(0.211)
D v. SP	0.431***	0.000	0.500***	0.002	0.907***	0.087
	(0.091)	(0.063)	(0.095)	(0.061)	(0.189)	(0.220)
SP v. D	0.131	0.010	0.184	0.048	0.269	0.169
	(0.092)	(0.063)	(0.096)	(0.061)	(0.191)	(0.219)
D v. D	0.067	0.030	0.093	0.032	0.175	0.242
	(0.091)	(0.063)	(0.095)	(0.061)	(0.188)	(0.216)
Constant	1.989***	1.359***	1.952***	1.293***		
	(0.066)	(0.045)	(0.068)	(0.043)		
cut1						
Constant					-0.648***	0.972***
					(0.137)	(0.160)
cut2						
Constant					0.939***	2.984***
					(0.139)	(0.182)
cut3						
Constant					3.003***	4.740***
					(0.163)	(0.272)
R-squared	0.028***	0.005	0.032***	0.008		
N	1709	1834	1538	1656	1709	1834

The base category is silence v. silence.

† OLS estimation, eliminating respondents who spent less than 30 seconds on the main experiment page.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3
Experiment 2: US

	OLS		Attentives†		Ologit	
	(1)	(2)	(3)	(4)	(5)	(6)
	CHN	JPN	CHN	JPN	CHN	JPN
China v. Japan						
SP v. S	0.191	0.084	0.285**	0.135	0.359	0.137
	(0.105)	(0.103)	(0.104)	(0.110)	(0.195)	(0.199)
S v. SP	-0.160	0.206*	-0.111	0.279**	-0.313	0.363
	(0.105)	(0.101)	(0.105)	(0.108)	(0.198)	(0.196)
SP v. SP	0.114	0.153	0.176	0.177	0.202	0.314
	(0.108)	(0.104)	(0.109)	(0.112)	(0.203)	(0.204)
D v. S	0.260*	0.038	0.347**	0.069	0.481*	0.087
	(0.109)	(0.107)	(0.110)	(0.115)	(0.204)	(0.209)
S v. D	0.037	0.349***	0.040	0.437***	-0.009	0.750***
	(0.107)	(0.104)	(0.107)	(0.111)	(0.204)	(0.206)
D v. SP	0.075	0.094	0.180	0.153	0.136	0.188
	(0.105)	(0.101)	(0.105)	(0.108)	(0.197)	(0.199)
SP v. D	-0.017	0.235*	0.042	0.266*	-0.078	0.448*
	(0.107)	(0.103)	(0.107)	(0.110)	(0.202)	(0.201)
D v. D	0.058	0.030	0.160	0.087	0.083	0.087
	(0.110)	(0.106)	(0.110)	(0.114)	(0.206)	(0.209)
Constant	1.975***	2.712***	1.801***	2.633***		
	(0.076)	(0.073)	(0.077)	(0.079)		
cut1						
Constant					-0.477***	-1.754***
					(0.145)	(0.155)
cut2						
Constant					0.906***	-0.616***
					(0.146)	(0.145)
cut3						
Constant					2.441***	1.327***
					(0.164)	(0.149)
R-squared	0.014**	0.013*	0.023***	0.017**		
N	1494	1491	1294	1292	1494	1491

The base category is silence v. silence.

† OLS estimation, eliminating respondents who spent less than 30 seconds on the main experiment page.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note on the frequencies of scrambling

The promotion of "show of force ($HiAct = 7$)" among *militerized interstate disputes* (MIDs) with low hostility level ("threat to use force" or "display of force" ($HostLev = 2$ or 3)) is calculated, using the Militarized Interstate Disputes (v4.1) dataset of the Correlates of War Project (Palmer et al., 2015). For example, MID number 4061 records aircraft scrambles between Japan and China in 1995, coding it "show of force." Of course, scrambles are just one type of showing force.

Balance tables

Starts from the next page.

Table 4
Experiment 1: Balance table: full sample

	Groups (1)		(2)		(3)		(4)		(5)		(6)		Total	
	%	CI	%	CI	%	CI	%	CI	%	CI	%	CI	%	CI
Age														
20-29 (n=279)	12	[9,16]	17	[13,22]	15	[11,19]	13	[10,17]	13	[10,17]	12	[9,16]	14	[12,15]
30-39 (n=454)	22	[18,26]	19	[15,24]	22	[18,27]	25	[21,30]	21	[17,26]	23	[19,28]	22	[20,24]
40-49 (n=466)	23	[19,28]	26	[21,31]	25	[21,30]	22	[18,27]	22	[18,27]	18	[14,22]	23	[21,24]
50-59 (n=407)	22	[18,26]	19	[15,23]	18	[14,22]	19	[15,24]	19	[15,23]	23	[19,28]	20	[18,22]
60-70 (n=453)	22	[18,26]	19	[15,24]	21	[17,25]	21	[17,26]	25	[21,30]	24	[20,29]	22	[20,24]
Total (n=2,059)	100		100		100		100		100		100		100	
Pearson: Uncorrected chi2(20) =	19.5438													
Design-based F(20.00, 41160.00) =	0.9767	Pr	=	0.49										
Gender														
Male (n=1,001)	50	[45,55]	47	[41,52]	46	[41,51]	47	[42,52]	51	[46,56]	50	[45,56]	49	[46,51]
Female (n=1,058)	50	[45,55]	53	[48,59]	54	[49,59]	53	[48,58]	49	[44,54]	50	[44,55]	51	[49,54]
Total (n=2,059)	100		100		100		100		100		100		100	
Pearson: Uncorrected chi2(5) =	3.2362													
Design-based F(5.00, 10290.00) =	0.6469	Pr	=	0.66										
Education														
Not college graduate (n=977)	48	[43,54]	48	[42,54]	45	[40,51]	48	[43,54]	48	[43,53]	47	[41,52]	47	[45,50]
College graduate (n=1,082)	52	[46,57]	52	[46,58]	55	[49,60]	52	[46,57]	52	[47,57]	53	[48,59]	53	[50,55]
Total (n=2,059)	100		100		100		100		100		100		100	
Pearson: Uncorrected chi2(5) =	1.0719													
Design-based F(5.00, 10290.00) =	0.2143	Pr	=	0.96										
Income (mil yen)														
less than 3.5 (n=501)	22	[18,27]	26	[21,31]	24	[20,29]	23	[19,28]	29	[25,34]	23	[19,27]	25	[23,26]
3.5-4.8 (n=331)	15	[12,19]	16	[12,20]	16	[13,20]	17	[13,21]	15	[12,20]	18	[14,23]	16	[15,18]
4.8-6.3 (n=408)	20	[16,24]	18	[14,23]	19	[15,23]	18	[14,22]	19	[15,24]	26	[21,31]	20	[18,22]
6.3-8.3 (n=334)	19	[15,23]	18	[14,23]	17	[14,22]	16	[13,21]	14	[11,18]	13	[10,17]	16	[15,18]
8.3- (n=470)	24	[20,29]	22	[18,27]	24	[20,28]	26	[21,31]	22	[18,27]	20	[16,25]	23	[21,25]
Total (n=2,044)	100		100		100		100		100		100		100	
Pearson: Uncorrected chi2(20) =	21.3637													
Design-based F(20.00, 40860.00) =	1.0677	Pr	=	0.38										

Table 5
Experiment 1: Balance table: full sample, cont.

	Groups													
	(1)		**(2)**		**(3)**		**(4)**		**(5)**		**(6)**		**Total**	
	%	CI	%	CI	%	CI	%	CI	%	CI	%	CI	%	CI
Ideology														
Ideology														
Liberal (n=116)	6	[4,9]	7	[4,11]	6	[4,9]	5	[3,8]	7	[5,10]	7	[5,10]	6	[5,8]
Moderate (n=1,358)	78	[73,82]	72	[66,77]	75	[70,80]	76	[70,80]	75	[70,79]	70	[64,74]	74	[72,76]
Conservative (n=357)	16	[13,21]	21	[17,27]	19	[15,23]	19	[15,24]	18	[14,23]	23	[19,28]	19	[18,21]
Total (n=1,831)	100		100		100		100		100		100		100	
Pearson: Uncorrected chi2(10) =	8.3327													
Design-based F(10.00, 18300.00) =	0.8328	Pr	=	0.60										
Region														
Hokkaido&Tohoku (n=225)	11	[8,15]	8	[6,12]	11	[8,15]	13	[10,17]	10	[8,14]	12	[9,16]	11	[10,12]
Kanto&Koshinetsu (n=870)	44	[39,50]	43	[38,49]	39	[34,44]	40	[35,45]	48	[42,53]	39	[34,44]	42	[40,44]
Chubu (n=319)	15	[12,19]	17	[13,21]	17	[14,21]	16	[12,20]	12	[9,16]	16	[13,21]	15	[14,17]
Kinki (n=355)	16	[12,20]	17	[13,22]	18	[14,22]	18	[14,22]	18	[14,23]	16	[13,21]	17	[16,19]
Chugoku&Shikoku (n=149)	7	[5,10]	8	[5,11]	8	[6,12]	7	[5,10]	5	[3,8]	9	[6,12]	7	[6,8]
Kyusyu&Okinawa (n=141)	6	[4,9]	7	[5,10]	6	[4,10]	7	[5,11]	7	[5,10]	7	[5,10]	7	[6,8]
Total (n=2,059)	100		100		100		100		100		100		100	
Pearson: Uncorrected chi2(25) =	18.0374													
Design-based F(25.00, 51450.00) =	0.7211	Pr	=	0.84										

Table 6
Experiment 1: Balance table: sample without DK

	Groups (1)		(2)		(3)		(4)		(5)		(6)		Total	
	%	CI	%	CI	%	CI	%	CI	%	CI	%	CI	%	CI
Age														
20-29 (n=217)	14	[10,19]	18	[14,23]	14	[10,20]	11	[8,15]	13	[10,18]	16	[11,21]	14	[13,16]
30-39 (n=322)	22	[18,28]	19	[15,25]	21	[16,27]	25	[21,31]	21	[17,26]	17	[13,23]	21	[19,23]
40-49 (n=329)	20	[15,25]	26	[21,32]	23	[18,29]	22	[18,27]	19	[15,24]	19	[14,25]	22	[20,24]
50-59 (n=311)	25	[20,31]	18	[14,23]	17	[12,22]	20	[16,25]	20	[16,25]	24	[19,31]	20	[18,23]
60-70 (n=342)	19	[15,25]	19	[14,24]	25	[20,32]	22	[17,27]	27	[22,32]	24	[19,30]	22	[20,25]
Total (n=1,521)	100		100		100		100		100		100		100	
Pearson: Uncorrected chi2(20) =	36.0525													
Design-based F(20.00, 41160.00) =	1.3310	Pr	=	0.15										
Gender														
Male (n=734)	49	[43,55]	47	[41,53]	44	[38,51]	46	[41,52]	50	[44,56]	53	[46,59]	48	[46,51]
Female (n=787)	51	[45,57]	53	[47,59]	56	[49,62]	54	[48,59]	50	[44,56]	47	[41,54]	52	[49,54]
Total (n=1,521)	100		100		100		100		100		100		100	
Pearson: Uncorrected chi2(5) =	5.2517													
Design-based F(5.00, 10290.00) =	0.7755	Pr	=	0.57										
Education														
Not college graduate (n=736)	50	[43,56]	48	[42,54]	46	[39,53]	47	[42,53]	50	[44,56]	50	[43,57]	48	[46,51]
College graduate (n=785)	50	[44,57]	52	[46,58]	54	[47,61]	53	[47,58]	50	[44,56]	50	[43,57]	52	[49,54]
Total (n=1,521)	100		100		100		100		100		100		100	
Pearson: Uncorrected chi2(5) =	1.8718													
Design-based F(5.00, 10290.00) =	0.2764	Pr	=	0.93										
Income (mil yen)														
less than 3.5 (n=369)	19	[14,24]	25	[20,31]	25	[19,31]	24	[19,29]	29	[25,35]	24	[18,30]	24	[22,27]
3.5-4.8 (n=238)	15	[11,20]	16	[12,21]	15	[11,21]	17	[13,21]	16	[12,20]	16	[12,22]	16	[14,18]
4.8-6.3 (n=295)	22	[17,27]	18	[14,23]	20	[15,26]	16	[12,21]	20	[16,25]	23	[18,29]	20	[18,22]
6.3-8.3 (n=247)	21	[16,26]	18	[14,23]	18	[13,23]	17	[13,22]	12	[9,16]	14	[10,19]	16	[15,18]
8.3- (n=360)	24	[19,30]	22	[18,28]	23	[18,29]	27	[22,32]	23	[19,28]	24	[18,30]	24	[22,26]
Total (n=1,509)	100		100		100		100		100		100		100	
Pearson: Uncorrected chi2(20) =	27.4454													
Design-based F(20.00, 40920.00) =	1.0111	Pr	=	0.44										

Table 7
Experiment 1: Balance table: sample without DK, cont.

	Groups													
	(1)		**(2)**		**(3)**		**(4)**		**(5)**		**(6)**		**Total**	
	%	CI	%	CI	%	CI	%	CI	%	CI	%	CI	%	CI
Liberal (n=94)	7	[4,11]	8	[5,12]	7	[4,11]	5	[3,9]	5	[3,9]	10	[6,15]	7	[6,8]
Moderate (n=1,031)	77	[71,82]	70	[64,75]	76	[70,82]	75	[69,80]	77	[71,81]	67	[60,73]	74	[71,76]
Conservative (n=273)	16	[12,22]	22	[18,28]	17	[12,23]	20	[16,25]	18	[14,23]	24	[18,30]	20	[18,22]
Total (n=1,398)	100		100		100		100		100		100		100	
Pearson: Uncorrected chi2(10) =	16.9266													
Design-based F(10.00, 19350.00) =	1.2216	Pr	=	0.27										
Region														
Hokkaido&Tohoku (n=161)	10	[7,14]	8	[6,12]	10	[7,15]	13	[9,17]	10	[7,14]	12	[9,18]	11	[9,12]
Kanto&Koshinetsu (n=652)	46	[39,52]	43	[37,49]	41	[34,47]	39	[34,45]	48	[42,53]	40	[33,47]	43	[40,45]
Chubu (n=224)	17	[13,22]	17	[13,22]	13	[9,18]	16	[12,21]	11	[8,15]	14	[10,20]	15	[13,17]
Kinki (n=267)	15	[11,20]	18	[14,23]	20	[15,26]	17	[13,21]	19	[15,24]	18	[13,23]	18	[16,20]
Chugoku&Shikoku (n=115)	8	[5,12]	8	[5,12]	10	[7,15]	8	[5,11]	5	[3,8]	7	[4,12]	8	[6,9]
Kyusyu&Okinawa (n=102)	5	[3,8]	6	[3,9]	6	[4,10]	8	[5,12]	7	[5,11]	9	[5,13]	7	[6,8]
Total (n=1,521)	100		100		100		100		100		100		100	
Pearson: Uncorrected chi2(25) =	30.6489													
Design-based F(25.00, 51450.00) =	0.9052	Pr	=	0.60										

Table 8
Experiment 2, ROK: balance table

	Groups									
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**Total**
	N	N	N	N	N	N	N	N	N	N
Age										
under 20 (n=15)	0	1	2	1	0	1	0	0	0	1
20-29 (n=411)	20	23	21	19	16	25	23	21	18	21
30-39 (n=397)	17	22	21	23	22	18	18	17	24	20
40-49 (n=467)	25	20	23	23	26	27	20	26	23	24
50-59 (n=481)	25	25	23	25	26	20	25	25	24	24
60-70 (n=212)	12	9	10	10	10	9	14	11	12	11
Total (n=1,983)	100	100	100	100	100	100	100	100	100	100
Pearson: Uncorrected chi2(40) =	30.6767									
Design-based F(40.00, 79279.90) =	0.7667	Pr =	0.86							
Gender										
Male (n=1,024)	52	55	53	51	59	49	49	47	51	52
Female (n=959)	48	45	47	49	41	51	51	53	49	48
Total (n=1,983)	100	100	100	100	100	100	100	100	100	100
Pearson: Uncorrected chi2(8) =	9.1173									
Design-based F(8.00, 15856.00) =	1.1391	Pr =	0.33							
Education										
Not college graduate (n=548)	29	29	24	28	24	29	31	31	24	28
College graduate (n=1,435)	71	71	76	72	76	71	69	69	76	72
Total (n=1,983)	100	100	100	100	100	100	100	100	100	100
Pearson: Uncorrected chi2(8) =	7.4357									
Design-based F(8.00, 15856.00) =	0.9290	Pr =	0.49							

Table 9
Experiment 2, ROK: balance table, cont.

	Groups									
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**Total**
	N	N	N	N	N	N	N	N	N	N
Income (mil won)										
under 2400 (n=266)	17	14	16	13	14	10	19	13	12	14
2400 - 4800 (n=700)	34	38	36	33	35	41	36	42	37	37
4800 and more (n=918)	49	48	48	54	50	49	46	45	50	49
Total (n=1,884)	100	100	100	100	100	100	100	100	100	100
Pearson: Uncorrected chi2(16) =	15.4696									
Design-based F(16.00, 30128.00) =	0.9663	Pr =	0.49							
Ideology										
Liberal (n=145)	8	9	9	7	8	6	8	7	9	8
Moderate (n=1,390)	77	74	75	77	76	77	75	77	74	76
Conservative (n=303)	16	17	16	16	16	18	16	17	17	16
Total (n=1,838)	100	100	100	100	100	100	100	100	100	100
Pearson: Uncorrected chi2(16) =	4.3070									
Design-based F(16.00, 29392.00) =	0.2690	Pr =	1.00							

Table 10
Experiment 2, US: balance table

	Groups									
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**Total**
	N	N	N	N	N	N	N	N	N	N
Age										
under 30 (n=469)	26	24	25	23	30	16	23	22	24	24
30-39 (n=398)	19	20	17	19	19	25	15	22	25	20
40-49 (n=388)	19	19	19	19	18	23	24	19	17	20
50-59 (n=431)	21	22	21	22	21	21	25	22	22	22
60-70 (n=299)	15	16	18	17	13	16	13	16	13	15
Total (n=1,985)	100	100	100	100	100	100	100	100	100	100
Pearson: Uncorrected chi2(32) =	28.6087									
Design-based F(32.00, 63488.00) =	0.8936	Pr =	0.64							
Gender										
Male (n=988)	51	54	49	51	43	51	48	50	49	50
Female (n=997)	49	46	51	49	57	49	52	50	51	50
Total (n=1,985)	100	100	100	100	100	100	100	100	100	100
Pearson: Uncorrected chi2(8) =	6.6649									
Design-based F(8.00, 15872.00) =	0.8327	Pr =	0.57							
Education										
Not college graduate (n=914)	48	43	43	47	47	42	48	51	45	46
College graduate (n=1,071)	52	57	57	53	53	58	52	49	55	54
Total (n=1,985)	100	100	100	100	100	100	100	100	100	100
Pearson: Uncorrected chi2(8) =	6.0318									
Design-based F(8.00, 15872.00) =	0.7536	Pr =	0.64							

Table 11
Experiment 2, US: balance table, cont.

	Groups									
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**Total**
	N	N	N	N	N	N	N	N	N	N
Income (thousand dollar)										
under 26 (n=375)	16	20	15	22	24	18	22	22	18	20
26 - 78 (n=923)	56	44	49	45	48	46	49	51	50	49
78 and more (n=595)	29	36	36	32	28	36	29	26	32	31
Total (n=1,893)	100	100	100	100	100	100	100	100	100	100
Pearson: Uncorrected chi2(16) =	21.1023									
Design-based F(16.00, 30272.00) =	1.3182	Pr =	0.18							
Ideology										
Liberal (n=318)	16	21	13	20	11	15	23	18	14	17
Moderate (n=999)	54	50	50	50	62	57	47	53	58	53
Conservative (n=560)	30	29	36	30	27	28	30	29	28	30
Total (n=1,877)	100	100	100	100	100	100	100	100	100	100
Pearson: Uncorrected chi2(16) =	25.9739									
Design-based F(16.00, 30016.00) =	1.6225	Pr =	0.06							

References

Beatty, P., & Herrmann, D. (1995). A framework for evaluating, "don't know" responses in surveys. In *Proceedings of the survey research methods section of the american statistical association* (pp. 1005–1010).

Berinsky, A. J., Margolis, M. F., & Sances, M. W. (2014). Separating the shirkers from the workers? Making sure respondents pay attention on self-administered surveys. *American Journal of Political Science*, *58*(3), 739–753.

Gilljam, M., & Granberg, D. (1993). Should we take don't know for an answer? *Public Opinion Quarterly*, *57*(3), 348–357.

Hainmueller, J., & Xu, Y. (2013). Ebalance: A stata package for entropy balancing. *Journal of Statistical Software*, *54*(7), 1–18.

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, *5*(3), 213–236.

Krosnick, J. A., Holbrook, A. L., Berent, M. K., Carson, R. T., Hanemann, W. M., Kopp, R. J., ... others (2002). The impact of "no opinion" response options on data quality: Non-attitude reduction or an invitation to satisfice? *Public Opinion Quarterly*, *66*(3), 371–403.

Palmer, G., d'Orazio, V., Kenwick, M., & Lane, M. (2015). The MID4 dataset, 2002–2010: Procedures, coding rules and description. *Conflict Management and Peace Science*, *32*(2), 222–242.

Statistics Bureau, Ministry of Internal Affairs and Communications. (2014). Table 5: Population by sex for prefectures - total population, japanese population. *Annual report of population estimates*. Retrieved from `http://www.e-stat.go.jp/SG1/estat/ListE.do?lid=000001132435`.

The Comparative Study of Electoral Systems. (2015). *CSES Module 4 second advance release [dataset].* doi:10.7804/cses.module4.2015-03-20.