Supplementary Material for

Reducing Item Non-Response to Vote Choice Questions: Evidence from a Survey Experiment in Mexico

Mollie J. Cohen
University of Georgia
Department of International Affairs
mj.cohen@uga.edu

Kaitlen J. Cassell
YouGov
kait.cassell@gmail.com

Table of Contents

A. Variable Coding

Table A1. Question wording summary

Variable Label	Question Wording
Experimental Treatments	
PREMXVB3N3- PREMXVB3N6; PREMXFUTBL *Vote Choice/ Soccer Experiment Preamble*	Now I am going to request that for the next question, you put these headphones on. I am going to place these clean covers over them for hygiene reasons, which we change for each interview. [INTERVIEWER INSTRUCTION: hand headset to respondent and connect to the handheld device] Before proceeding with the question, we'll test the functioning of the headphones. Please listen to the following audio and indicate if you can hear the audio well or if you need the volume adjusted. [INTERVIEWER INSTRUCTION: press the green "view/ play" button. Wait for the respondent to indicate that everything works well, make any necessary changes to volume, and then continue to the next question] (1) Everything works as it should (2) The person did not want to participate in the use of audio [respondent reassigned at random to C1 or C2]
MXVB3N1 *Vote Choice – Control*	You will hear the following question with response options numbered from 1 to 5 and then you will indicate the **number** that corresponds to your response. Who did you vote for as President in the last presidential elections of 2018? NOTE: response order and numbers randomized (1) None (went to vote but left the ballot blank or spoiled it) (2) Andrés Manuel López Obrador (MORENA) (3) Ricardo Anaya (PAN) (4) José Antonio Meade (PRI) (5) Other [always the final option]
MXVB3N2- MXVB3N6 *Vote Choice – Treatments*	Before continuing to the next question, I want to [Confidentiality reminder: remind you that all of the responses

	you provide will be kept confidential and anonymous] [**Anonymity guarantee: alert you that I will not know what response you give to the question, because, as you can see, my screen only shows the numbers of the options and the order in which responses appear changes in each interview**].
	You will hear the following question with response options numbered from 1 to 5 and then you will indicate the **number** that corresponds to your response. [*Audio conditions: Ask the interviewer to repeat the audio if necessary.*]
	Who did you vote for as President in the last presidential elections of 2018?
	NOTE: response order and numbers randomized
	(2) None (went to vote but left the ballot blank or spoiled it) (2) Andrés Manuel López Obrador (MORENA) (3) Ricardo Anaya (PAN) (4) José Antonio Meade (PRI) (5) Other [always the final option]
MXFUTBL1-MXFUTBL2 *Soccer Treatment*	You will hear the following question with response options numbered from 1 to 5 and then you will indicate the **number** that corresponds to your response.
	Which of the following players from the national soccer team do you think played the best in the World Cup in Russia?
	NOTE: response order and numbers randomized
	(1) Giovanni Dos Santos (2) Javier (el "Chicharito") Hernández (3) Hirving Lozano (4) Héctor Herrera (5) Other [always the final option]

Additional Variables

MXSSAT1G/ MXSSAT1H *Enjoyment of survey*	How much did you enjoy participating in this survey? (1) A lot (2) Somewhat (3) A little (4) Not at all

	For all analyses in paper, this variable has been recoded so that higher values indicate greater enjoyment. The order of MXSSAT1 and MXSSAT2 were varied at random across surveys.
MXSSAT2G/ MXSSAT2H *Enjoyment of audio question*	And now thinking about the question you received by audio in comparison to the other questions on the survey. Did you enjoy this question… (1) a lot more (2) more (3) the same (4) less, or (5) much less than the other questions? *For all analyses in paper, this variable has been recoded so that higher values indicate greater enjoyment. The order of MXSSAT1 and MXSSAT2 were varied at random across surveys.*
MXCHECKVB3 *Anonymity Beliefs*	How much do you believe that your response to the question about your vote in the 2018 presidential election was confidential, that is, that I don't know what response you gave? (1) A lot (2) Somewhat (3) A little (4) Not at all *For all analyses in paper, this variable has been recoded so that higher values indicate greater enjoyment.*

Control Variables

Edad *Age*	Respondent's age by cohort: 18-25, 26-35, 36-45, 46-55, 56-65, 66+
Mujer *Gender*	Dummy variable: male = 0, female = 1
Edr *Education Level*	Education level respondent reports completing (None, Primary, Secondary, Post-Secondary)
i_o4 *Income*	Household monthly income, measured by level (none, less than $620; $620-$1,200; $1,201-$1,600; $1,601-$2,000; $2,001-$2,400; $2,401-$2,800; $2,801-$3,400; $3,401-$3,800;

	$3,801-$4,300; $4,301-$4,800; $4,801-$5,400; $5,401-$6,400; $6,401-$7,600; $7,601-$9,000; $9,001-$13,000; more than $13,000)
Estado *State*	Mexico City = 1001, Mexico State=1002

B. Balance Checks

Table B1 presents the results of a multinomial logistic regression examining differences across groups assigned to each treatment condition. Because we observe significant differences in education, some age categories, and residence in Mexico State vs. Mexico City, we control for these factors in robustness checks presented in Models 4-6 in Table C1 below.

Table B1. Balance Checks – Vote Treatment (Multinomial Logit)

	Confidentiality Reminder	Audio, No Reminder	Audio + C. Reminder	Audio + A. Guarantee	Audio + A. and C.
Woman	0.133	-0.116	-0.247	0.039	-0.147
	(0.232)	(0.207)	(0.218)	(0.209)	(0.204)
Household Income	-0.003	-0.001	0.005	-0.009	-0.006
	(0.022)	(0.025)	(0.020)	(0.018)	(0.022)
Education					
Primary Education	-1.925	-2.054	-2.387	-1.517	-2.220
	(1.097)	(1.137)	(1.093)	(1.188)	(1.110)
Secondary Education	-1.649	-1.581	-1.938	-1.544	-1.471
	(1.098)	(1.142)	(1.131)	(1.231)	(1.114)
Post-Secondary Education	-1.337	-1.242	-1.660	-0.691	-0.814
	(1.129)	(1.162)	(1.181)	(1.265)	(1.111)
Age					
26-35	-0.033	-0.354	-0.030	-0.736	0.047
	(0.358)	(0.341)	(0.347)	(0.334)	(0.314)
36-45	-0.247	-0.576	-0.582	-0.217	-0.152
	(0.356)	(0.355)	(0.357)	(0.335)	(0.313)
46-55	-0.240	-0.199	-0.2-2	-0.117	0.210
	(0.383)	(0.332)	(0.361)	(0.353)	(0.326)
56-65	0.357	0.204	0.602	0.033	0.350
	(0.444)	(0.401)	(0.419)	(0.454)	(0.391)
66+	-0.153	-0.045	-0.232	-0.551	0.692
	(0.485)	(0.380)	(0.485)	(0.461)	(0.402)
Mexico State	0.160	0.276	-0.112	0.269	0.424
	(0.181)	(0.172)	(0.214)	(0.192)	(0.211)
Constant	1.555	1.687	2.065	1.346	1.088
	(1.156)	(1.173)	(1.093)	(1.188)	(1.184)
Observations	1,201	1,201	1,201	1,201	1,201

Survey-adjusted standard errors in parentheses. Results estimated without adjustments for complex survey design yield substantively similar results. Missing observations are mostly attributable to missing values for household income.

Table B1. Balance Checks – Soccer Treatment (Logit)

	Audio Soccer Treatment
Woman	0.188
	(0.119)
Household Income	0.001
	(0.013)
Education	
Primary Education	0.819
	(0.470)
Secondary Education	0.540
	(0.467)
Post-Secondary Education	0.175
	(0.514)
Age	
26-35	0.235
	(0.191)
36-45	0.256
	(0.201)
46-55	-0.036
	(0.215)
56-65	-0.127
	(0.234)
66+	-0.057
	(0.228)
Mexico State	-0.135
	(0.120)
Constant	-1.212
	(0.453)
Observations	1201

Survey-adjusted standard errors in parentheses. Results estimated without adjustments for complex survey design yield substantively similar results. Missing observations are mostly attributable to missing values for household income.

C. Complete Results and Robustness Checks

This section provides complete results and robustness checks for tests of the four core hypotheses presented in the paper body.

Table C1. Logistic Regression Models with Adjusted Standard Errors, Controls (Corresponds to Figure 1)

	Model 1. Non-Response	Model 2. Non-Response	Model 3. Non-Response	Model 4. Non-Response	Model 5. Non-Response	Model 6. Non-Response
Confidentiality Reminder	0.027	0.027	0.023	0.038	0.038	0.035
	(0.326)	(0.263)	(0.325)	(0.328)	(0.248)	(0.329)
Audio, No Reminder	-0.368	-0.368	-0.372	-0.393	-0.393	-0.397
	(0.354)	(0.369)	(0.363)	(0.356)	(0.386)	(0.364)
Audio, C. Reminder	-0.571	-0.571	-0.571	-0.588	-0.588	-0.586
	(0.387)	(0.383)	(0.362)	(0.390)	(0.368)	(0.358)
Audio, A. Guarantee	-1.126	-1.126	-1.126	-1.123	-1.123	-1.118
	(0.475)	(0.510)	(0.448)	(0.477)	(0.502)	(0.456)
Audio, A&C Guarantee	-0.529	-0.529	-0.529	-0.668	-0.668	-0.670
	(0.387)	(0.364)	(0.359)	(0.403)	(0.409)	(0.365)
Age				0.259	0.259	0.260
				(0.078)	(0.086)	(0.074)
Education				0.049	0.049	0.050
				(0.169)	(0.190)	(0.159)
Mexico State				-0.116	-0.116	-0.110
				(0.232)	(0.227)	(0.211)
Constant	-2.429	-2.429	-2.425	112.819	112.819	107.050
	(0.233)	(0.255)	(0.218)	(232.242)	(227.550)	(211.778)
Unadjusted SEs	X			X		
Interviewer Clustered SEs		X			X	
Survey-Adjusted SEs			X			X
Observations	1,432	1,432	1,428	1,430	1,430	1,426

Columns present the results of logistic regression models estimating the effect of each variable on item non-response to the vote choice question. Standard errors are reported in parentheses.

Table C2. Predicted Probability of Item Non-Response (corresponds to Figure 2)

	Predicted Probability (SE)	Z-score (P-value)
Audio Treatment		
Audio	0.053	5.24
	(0.010)	(0.000)
Interviewer	0.082	6.68
	(0.012)	(0.000)
Difference	-0.029	-1.84
	(0.016)	(0.065)
N	993	
Confidentiality Reminder		
Confidentiality	0.066	5.85
	(0.011)	(0.000)
No Confidentiality	0.069	6.13
	(0.011)	(0.000)
Difference	-0.003	-0.18
	(0.016)	(0.855)
N	993	

Columns present the predicted probability of non-response following logistic regression models estimating the effect of each variable on item non-response to the vote choice question. Unadjusted standard errors are reported in parentheses.

Table C3. Logistic Regression, Predicting AMLO's Vote

	Mexico City		Mexico State	
	Model 1	Model 2	Model 3	Model 4
Confidentiality Reminder	0.063	0.066	-0.120	-0.155
	(0.295)	(0.296)	(0.269)	(0.271)
Audio, No Reminder	-0.087	-0.086	-0.291	-0.297
	(0.288)	(0.289)	(0.263)	(0.265)
Audio, C. Reminder	-0.601	-0.590	-0.586	-0.619
	(0.276)	(0.277)	(0.280)	(0.282)
Audio, A. Guarantee	-0.616	-0.566	-0.043	-0.074
	(0.289)	(0.291)	(0.276)	(0.278)
Audio, A&C Guarantee	-0.636	-0.607	-0.071	-0.126
	(0.295)	(0.297)	(0.270)	(0.272)
Age		-0.030		-0.039
		(0.054)		(0.056)
Education		-0.260		0.245
		(0.123)		(0.131)
Constant	0.898	1.524	0.525	0.206
	(0.207)	(0.411)	(0.195)	(0.401)
Observations	657	656	692	692

Columns present the results of logistic regression models estimating the effect of each variable on reporting a vote for AMLO. Unadjusted standard errors are reported in parentheses.

Table C4. Predicted Probability of Vote for AMLO (corresponds to Figure 3)

Mexico City				
	Predicted Probability (SE)	Chi-square (P-value)	Predicted Probability (SE)	Chi-square (P-value)
Control	0.711	9.88	0.709	9.56
	(0.042)	(0.002)	(0.043)	(0.002)
Confidentiality Reminder	0.723	11.96	0.723	11.75
	(0.042)	(0.001)	(0.042)	(0.001)
Audio, No Reminder	0.692	7.30	0.691	7.08
	(0.043)	(0.007)	(0.043)	(0.008)
Audio, C. Reminder	0.574	0.01	0.575	0.00
	(0.045)	(0.943)	(0.045)	(0.961)
Audio, A. Guarantee	0.570	0.02	0.581	0.01
	(0.050)	(0.888)	(0.050)	(0.941)
Audio, A&C Guarantee	0.565	0.05	0.571	0.01
	(0.052)	(0.820)	(0.052)	(0.904)
Observations	657		656	
Controls			X	
Mexico State				
	Predicted Probability (SE)	Chi-square (P-value)	Predicted Probability (SE)	Chi-square (P-value)
Control	0.628	3.44	0.635	4.02
	(0.045)	(0.064)	(0.045)	(0.045)
Confidentiality Reminder	0.600	1.57	0.599	1.47
	(0.045)	(0.211)	(0.045)	(0.226)
Audio, No Reminder	0.558	0.10	0.564	0.20
	(0.044)	(0.746)	(0.044)	(0.651)
Audio, C. Reminder	0.485	1.39	0.484	1.42
	(0.050)	(0.239)	(0.050)	(0.234)
Audio, A. Guarantee	0.618	2.56	0.618	2.51
	(0.046)	(0.109)	(0.047)	(0.113)
Audio, A&C Guarantee	0.612	2.33	0.606	1.88
	(0.044)	(0.127)	(0.045)	(0.170)
Observations	692		692	
Controls			X	

Columns present the predicted probability of voting for AMLO following logistic regression models estimating the effect of each variable on vote choice. Models estimated using unadjusted standard errors. Chi-square values and p-tests estimating the null hypothesis that the predicted probability is no different from observed vote rates: 0.577 in Mexico City and 0.544 in Mexico State. Significant coefficients are significantly different from AMLO's official vote share.

Table C5. Design Effects for Models Clustering by PSU, Interviewer (Coefficients in Table C1)

	Item non-response: DEFF DEFT		Item non-response: DEFF DEFT	
	Model 1	Model 4	Model 2	Model 5
Confidentiality Reminder	0.648	0.578	0.991	1.021
	0.805	0.761	0.996	1.011
Audio, No Reminder	1.087	1.178	1.050	1.050
	1.043	1.085	1.025	1.024
Audio, C. Reminder	0.978	0.891	0.871	0.841
	0.989	0.944	0.934	0.917
Audio, A. Guarantee	1.149	1.085	0.889	0.900
	1.072	1.042	0.943	0.949
Audio, A&C Guarantee	0.884	1.055	0.860	0.838
	0.940	1.027	0.927	0.916
Controls		X		X
Interviewer Clustered SEs	X	X		
Survey-Adjusted SEs			X	X
Observations	1,432	1,432	1,428	1,430

D. Who refused the audio treatments?

Of the 932 participants who were assigned an audio treatment condition, 77 individuals opted not to receive the treatment and were randomly assigned to either treatment 1 or 2, the non-audio conditions. We explore three possibilities: 1) that interviewers may have implicitly or explicitly encouraged individuals to refuse treatment; 2) that individuals who opted out of the audio treatment differ on observable features from individuals that were successfully treated; and 3) that individuals who opted out of the audio condition behaved differently with respect to their vote choice.

1) Interviewer effects

Some interviewers had more respondents choose to refuse the audio treatment. Table D1 shows audio treatment refusals (levels and rates) for both the vote choice and soccer items. Interviewers who had 10% or more of participants refuse either treatment are shaded in gray. A small number of interviewers stand out as having a significant number of respondents refuse audio treatments. For example, more than two-thirds of Interviewer 41's participants who were assigned to an audio treatment refused, for both the vote choice and soccer questions.

These results are consistent with interviewer effects, with some interviewers discouraging respondents from being treated, and perhaps shaping response to the survey item. On the other hand, interviewers' work areas are geographically assigned. Usually, interviewers are assigned to recruit participants from a particular neighborhood; individuals nested in neighborhoods tend to have similar levels of education and income, and to be of a similar age. Table D2 below shows that age is associated with audio refusal. It is also possible that some interviewers with high refusal rates were simply assigned to conduct interviews in neighborhoods where the likelihood of refusal was unusually high.

In addition to trends in audio refusal, also see wide variation in the number of surveys conducted by each interviewer. Whereas Interviewer 1 had 45 respondents assigned to treatment, and Interviewer 7 had 61 respondents assigned to one of the audio conditions, Interviewer 3 only had seven total respondents assigned to treatment. Due to the wide variation in the number of interviews assigned to each interviewer from each audio condition (and due also to the very small – effectively zero – inter-class correlation values we estimate when running such models), we choose not to run a hierarchical model including random slopes for interviewer effects. We do, however, cluster our standard errors by interviewer in robustness checks presented in Table C1 to account for any noise caused by these individuals. Doing so does not substantially change our results.

D1. Reassignment Totals and Proportions by Interviewer

Interviewer	Reassigned (vote)	Voters Assigned Treatment	%Reassigned (vote)	Reassigned (soccer)	Assigned Soccer Treatment	%Reassigned (soccer)
1	1	23	4.35%	0	22	0.00%
2	--	--		0	1	0.00%
3	0	4	0.00%	0	3	0.00%

4	0	6	0.00%	0	4	0.00%
5	0	3	0.00%	0	4	0.00%
6				1	3	33.33%
7	0	40	0.00%	0	21	0.00%
8	0	4	0.00%	0	8	0.00%
9	0	1	0.00%	--	--	--
11	5	37	13.51%	5	31	16.13%
12	0	4	0.00%	0	6	0.00%
13	2	18	11.11%	2	15	13.33%
				0	1	0.00%
15	4	24	16.67%	11	19	57.89%
16	0	6	0.00%	--	--	--
17	0	31	0.00%	0	24	0.00%
18	1	52	1.92%	3	45	6.67%
19	0	1	0.00%	--	--	--
20	0	1	0.00%	--	--	--
21	9	18	50.00%	10	23	43.48%
22	0	6	0.00%	1	4	25.00%
24	1	14	7.14%	0	15	0.00%
25	0	1	0.00%	--	--	--
27	0	5	0.00%	0	4	0.00%
28	2	23	8.70%	1	16	6.25%
29	2	45	4.44%	4	48	8.33%
30	0	52	0.00%	2	25	8.00%
31	0	1	0.00%	--	--	--
32	0	1	0.00%	0	2	0.00%
33	2	34	5.88%	0	35	0.00%
34	0	3	0.00%	0	2	0.00%
35	0	2	0.00%	0	1	0.00%
36	2	31	6.45%	2	22	9.09%
37	2	54	3.70%	4	43	9.30%
38	0	7	0.00%	0	5	0.00%
39	0	8	0.00%	0	4	0.00%
40	0	6	0.00%	0	4	0.00%
41	14	20	70.00%	8	12	66.67%
42	10	86	11.63%	5	55	9.09%
43	1	4	25.00%	0	3	0.00%
44	0	6	0.00%	0	4	0.00%
45	0	22	0.00%	0	10	0.00%
47	0	6	0.00%	1	9	11.11%

48	2	15	13.33%	14	33	42.42%
49	0	7	0.00%	0	10	0.00%
50	1	3	33.33%	0	1	0.00%
51	0	11	0.00%	0	8	0.00%
52	0	1	0.00%	0	6	0.00%
53	0	8	0.00%	0	6	0.00%
54	0	32	0.00%	0	15	0.00%
55	0	1	0.00%			
57	0	64	0.00%	2	49	4.08%
58	15	70	21.43%	11	47	23.40%
59	1	10	10.00%	0	4	0.00%
61	--	--	--	0	2	0.00%
62	--	--	--	0	1	0.00%
Total	77	932	8.26%	87	735	11.84%

2) Demographic differences

Another possibility is that individuals self-selected out of the audio condition based on some observable characteristic. We examine whether the education, age, or gender varies systematically across groups, using a series of student's t-tests. Because assignment to the vote choice question was limited to self-identified voters, and random assignment to the soccer question was blocked on voting, we only estimate results for voters here. Caution is warranted insofar as the sample size of individuals who opted out of the audio treatments are small – 77 individuals for the vote choice question and 58 individuals (of 87 total refusals) for the soccer question.

Table D2. Demographics of (Un)successfully Treated Respondents (voters only)

Variable	Mean – s*uccessfully treated* (SE)	Mean – *refused treatment* (SE)	T-Stat (p-value)
Vote Choice Question			
Age	43.0 (0.444)	51.8 (2.057)	-4.61 ($p < 0.01$)
Education	10.5 (0.116)	9.9 (0.561)	1.28 ($p = 0.20$)
Gender	47.7% male	45.5% male	-0.38 ($p = 0.70$)
Soccer Question			
Age	42.3 (0.763)	48.8 (2.314)	-2.87($p < 0.01$)
Education	10.1 (0.630)	10.3 (0.200)	-0.33 ($p = 0.74$)
Gender	44.4% male	39.7% male	-0.68 ($p = 0.50$)

For the vote choice question, we find significant differences in age: individuals who opted out of the audio treatment were older. In substantive terms, those who opted out of the audio treatment were about nine years older than those who were successfully treated. Similarly, for the soccer

treatment we only find a significant difference for age; refusers are about 6.5 years older. There is no significant difference in the gender makeup of the two groups for either question administered by audio.

3) Differences in item non-response

A final possibility that we considered is that the individuals who opted out of the audio treatment were also more (less) likely to answer the normal vote choice question. We examine this possibility by calculating a student's t-test.

Table D3. Item Non-Response to Vote Choice by Treatment Acceptance (groups 3-6)

	Non-response – *successfully treated*	Non-response – *refused treatment*	T-Stat (p-value)
Vote Choice	3.74%	14.3%	-4.26 (p = 0.00)

We find that individuals that opted out of the audio treatment are significantly less likely to respond to the vote choice question compared to individuals that were successfully treated. By analyzing these individuals according to their original group assignment, we are making it more difficult to uncover significant effects.

E. Mechanisms Tests

This section presents tests of the two mechanisms examined in the paper body: improved perceptions of response anonymity, and increased survey enjoyment and attention due to the novelty of the audio item.

Table E1. Beliefs about Vote Choice Response Anonymity by Treatment Condition (corresponds to Table 2)

	Anonymity Perceptions (fully crossed)	Anonymity Perceptions (all voters)	Anonymity Perceptions (audio only)	Anonymity Perceptions (all voters)
Audio Treatment – Vote	0.015	0.106		0.011
	(0.069)	(0.059)		(0.067)
Confidentiality Assurance			-0.095	-0.093
			(0.095)	(0.067)
Anonymity Guarantee			0.243	0.242
			(0.085)	(0.078)
Constant	2.992	2.992	3.051	3.039
	(0.049)	(0.048)	(0.065)	(0.059)
Observations	964	1396	914	1396

Table presents OLS estimates. Models estimated without covariates. Unadjusted standard errors are reported in parentheses.

Table E2. Item Non-Response to the Soccer Question by Condition

	Non-response: voters (SE)	Z-score (P-value)	Non-response: abstainers (SE)	Z-score (P-value)
Audio Treatment - Soccer	0.289	2.46	-0.012	-0.06
	(0.117)	(0.014)	(0.204)	(0.952)
Constant	-0.855	-11.94	-0.845	-5.79
	(0.072)	(0.000)	(0.146)	(0.000)
	Predicted Probability (SE)	T-statistic – difference (P-value)	Predicted Probability (SE)	T-statistic – difference (P-value)
Control	0.298	-2.467	0.300	0.060
	(0.015)	(0.014)	(0.031)	(0.952)
Audio Treatment - Soccer	0.362		0.298	
	(0.022)		(0.030)	
Observations	1,430		458	

Top panel of table presents results from a logistic regression model, and bottom panel presents predicted probabilities. Models estimated without covariates. Unadjusted standard errors are reported in parentheses.

Table E3. Survey Enjoyment by Condition (tests of H_{2m})

	Audio enjoyment (voters)	Audio enjoyment (voters)	Survey enjoyment (voters)	Survey enjoyment (abstainers)
Audio Treatment - Soccer	0.382	-0.398	-0.082	0.020
	(0.065)	(0.063)	(0.044)	(0.083)
Constant	2.869	-10.505	30.464	43.042
	(0.053)	(60.228)	(0.043)	(83.524)
Controls		X	X	X
Observations	1,271	1,269	1,423	455

Table presents logistic regression estimates. Unadjusted standard errors are reported in parentheses.

Table E4. Logit – Item Non-Response to the Soccer Question by Treatment, Gender

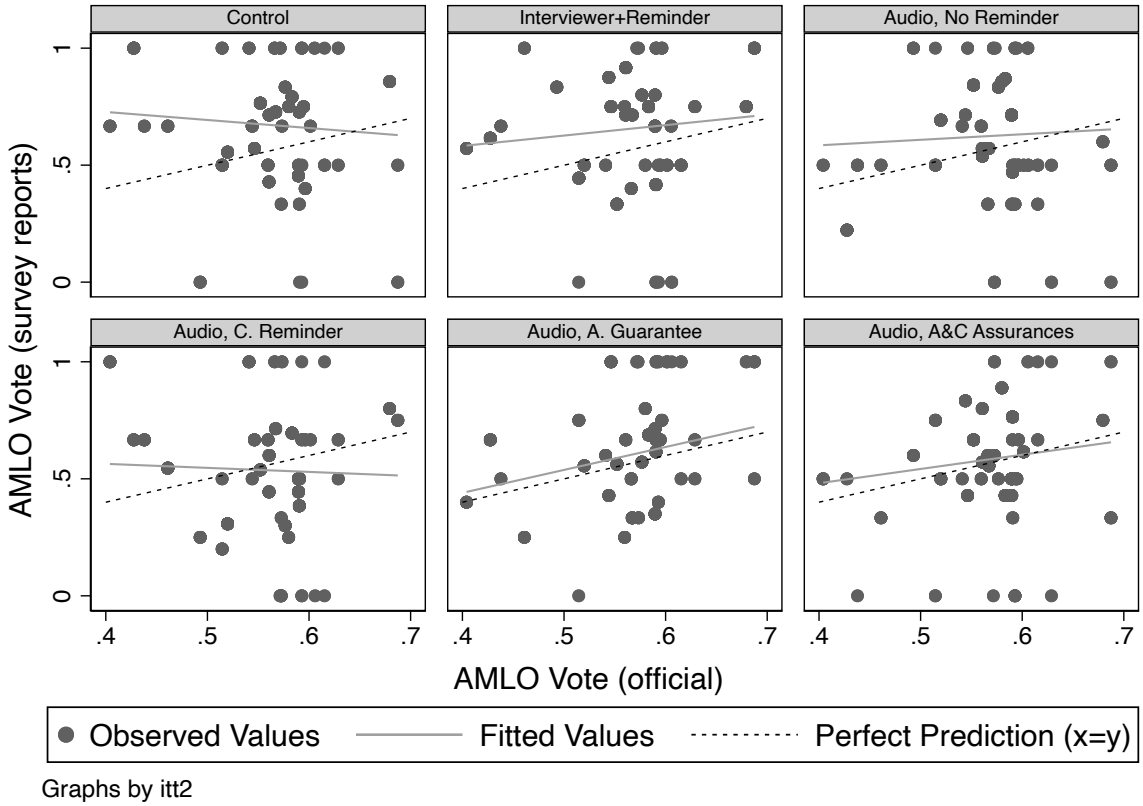
	Non-Response (voters)	Non-Response (abstainers)
Audio Treatment - Soccer	0.396	0.634
	(0.189)	(0.326)
Woman	0.857	1.211
	(0.148)	(0.316)
Audio Treatment - Soccer X Woman	-0.237	-1.108
	(0.243)	(0.426)
Constant	-1.325	-1.544
	(0.114)	(0.253)
Observations	1432	458

Table presents logistic regression estimates. Unadjusted standard errors are reported in parentheses. Models estimated without covariates.

F. Contextual Features and Item Non-Response

In addition to the two mechanisms we detail in the manuscript, we expected that context would influence our treatments' effectiveness. In particular, we expected that the returns from more private treatments would be higher for respondents living in very non-competitive locales – that is, where AMLO's vote share was very high or very low. Ultimately, the sample included very few such locales, making it challenging to confidently estimate the interaction between context and treatment. We therefore present scatterplots of AMLO's estimated vote share according to survey responses (y-axis) and his observed vote share according to INE (x-axis) by municipality. The solid gray line shows the line of best fit through these data points; the dashed line would represent perfect reporting. Scatterplots are separated by treatment condition.

The estimated lines in the audio condition with an anonymity guarantee and the audio condition with anonymity and confidentiality assurances most closely track perfect reporting. Ultimately, we have limited certainty around these estimates when we divide the data by treatment, so we leave these questions for future work.



G. Nicaragua Study – Description and Results[1]

As part of a nationally representative study (N=1,591) fielded in Nicaragua from August 16-September 21, 2017, respondents were asked about their voting behavior in the 2016 presidential election. Following a standard turnout question, the 873 individuals who reported voting in 2016 were assigned at random to one of three groups. 234 voters (26.8%) were assigned to a control group and asked to report their retrospective voting behavior using the following question:

> *Who did you vote for in the last presidential elections in 2016?*
> *(0) None, went to vote but left the ballot blank or spoiled it*
> *(502) Daniel Ortega, Frente Sandinista de Liberación Nacional – FSLN*
> *(506) Maximino Rodríguez, Partido Liberal Constitucionalista – PLC*
> *(577) Other candidate*

"Does not know" and "does not answer" were available as unread options. The order of the first three response options was assigned at random; the "other" option was anchored as the last item.[2]

The first treatment condition reminded study participants that their survey responses were anonymous. 228 (26.1%) respondents who indicated having voted in 2016 received the following reminder, drawn from the study information sheet, prior to receiving the vote question:

> *"Before moving on to the next question, I want to remind you that all of the answers that you provide will be kept confidentially and anonymously."*

Following this reminder, a vote question identical to that employed in the control group was asked. As in the control group, response options were listed in random order, with "other" anchored as the final option.

411 voters (47.1%) were assigned to a second treatment group. These individuals received the same vote question with randomly ordered response options following a privacy reminder similar to that in the Reminder only condition. The vote question was then administered via anonymized audio recording.[3] Following a test question to assure that the volume was correctly adjusted, respondents in the Audio condition were read the following script by the interviewer:

> *"You will hear the next question with response options numbered from 1 to 4 and then you will indicate to me the* ***number*** *that corresponds to your answer. Your answer to this question is* ***completely anonymous.*** *I will not know your response to the question since, as*

[1] Because the Nicaraguan electoral commission does not report complete disaggregated vote totals, it is not possible to know whether response accuracy improved or declined across conditions. The Mexico study was designed to improve on this issue, and to correct slight variations in the wording of anonymity guarantees across conditions.

[2] Several additional candidates appeared on the ballot but received less than 5% of the vote. To simplify the response task, these candidates were grouped into a single "other" category, which was always the final option. Because the "other" option included multiple candidates, giving this response did not reveal the voter's choice to the interviewer. Neither interviewers nor respondents were informed that the "other" category was anchored.

[3] Half of respondents in the audio group were randomly assigned to a recording with a female voice, and half received an identical recording with a male voice. I find no consistent differences by audio gender (including shared gender with the respondent), so group both audio conditions together here.

you can see, my screen only shows the numbered options, and the response order is assigned at random."[4]

Interviewers were then instructed to show the screen of the electronic device to the respondent, to verify that the numbers were not associated with responses. The interviewer saw only the question label, written instructions for reading the introductory script, a green "play" button, and a list of numbered radio buttons. A vote choice question identical to that described above was then administered via audio recording, with one modification: each response option was preceded by a number (e.g., "1. None, went to vote but left the ballot blank or spoiled it").

Table G.1 shows the individual elements of the question asked to each group. Respondents in the Control condition received the vote question without additional reminders or tasks to anonymize their responses. Those in the "Reminder" condition received only a privacy reminder, while those in the "Audio" group received a privacy reminder similar (though not identical) to that read in the reminder condition *and* a question that anonymized response options.

Table G1. Elements of Questions by Condition

	Control	Reminder	Audio
Privacy Reminder		X	X
Anonymized Response			X

Overall, the randomization worked as expected: groups were balanced across a range of sociodemographic covariates (gender, urban residence, region of residence, and education), although younger individuals were slightly more likely to receive the experimental treatments, and wealthier individuals were slightly more likely to be in the control condition. For quality control purposes, all vote questions were recorded; nearly all respondents in the audio condition provided numbered response options rather than names.

Item Non-Response

How did the treatment condition the likelihood of refusing to respond to the vote question? The dependent variable in this analysis takes the value of 1 if the respondent refused to respond to the vote choice question, and 0 if she provided a substantive response. "Don't know" responses are excluded.[5]

Figure G1 shows the likelihood that an average respondent from the Control, Reminder, and Audio conditions refuses to respond to the vote choice question.[6] The likelihood of non-response drops substantially when sensitive question techniques are employed. Compared to 16.2% non-response in the Control group, non-response drops 10.1 percentage points, to 6.1%, in the Reminder condition. Non-response drops even more in the Audio condition, to 3.5% - a

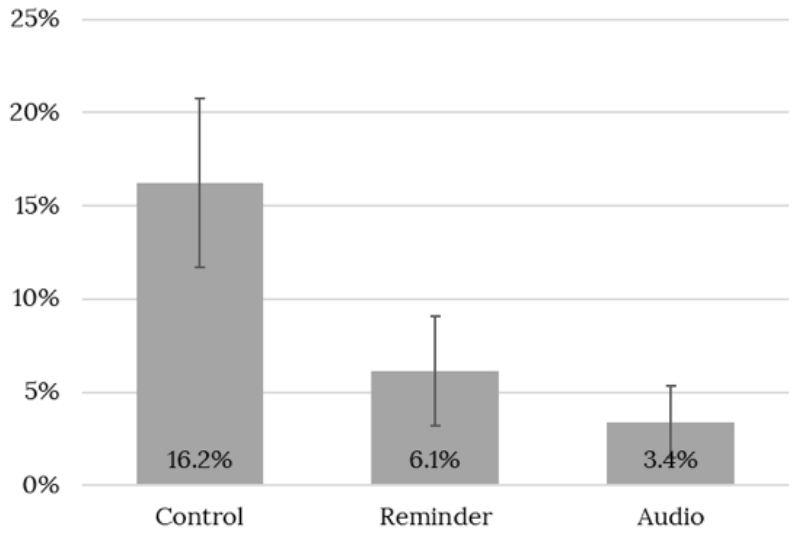
[4] The slightly stronger wording for the anonymity guarantee used in the Audio condition may have increased the credibility of the guarantee.

[5] The frequency of "don't know" responses was similar in the Control and Audio groups (1.7% and 1.22%, respectively), but about twice as common in the Reminder condition (3.5% of responses). Results for models including "don't know" responses are consistent with those shown here.

[6] No covariates are included in this model, and results from such models are substantively similar.

12.8 percentage point decrease. T-tests confirm that differences between the Control condition and both the Reminder and Audio groups are statistically significant with p<.01.

Figure G1. Non-Response to Vote Question by Condition



N=873. Bars represent non-response rates by condition and the vertical whiskers represent a 95% confidence interval around the estimate.

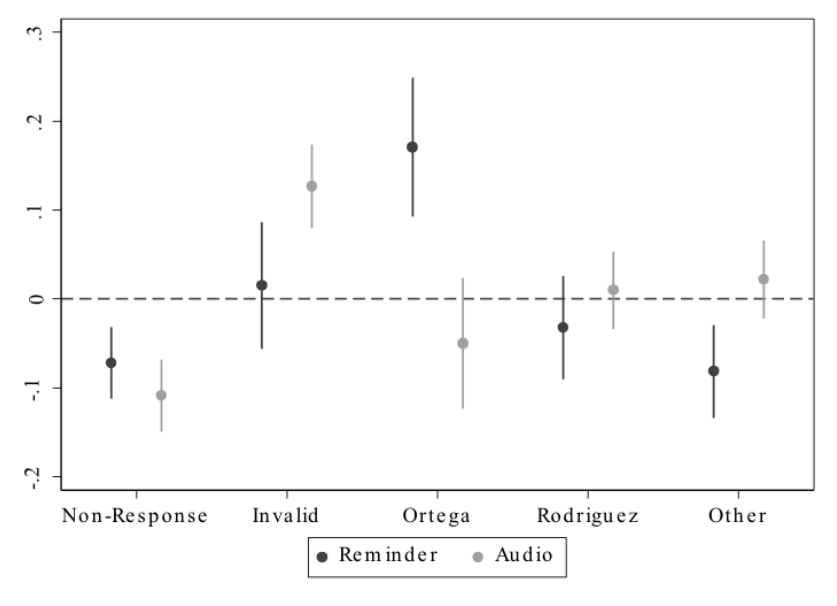
As expected, both the reminder and audio treatments substantially reduced non-response to the vote choice question. Although non-response was lower in the audio condition than in the reminder condition, these differences are not statistically significant.

Differences in Reported Vote Choice by Condition

The study also documented significant variation in the *content* of individuals' responses by treatment condition below. In particular, respondents in the Audio condition are more likely to report invalidating their ballots, consistent with literature suggesting that ACASI modes increase the reporting of sensitive behaviors. Unexpectedly, individuals in the Reminder condition are significantly more likely to report a vote for incumbent president Daniel Ortega; this trend is particularly pronounced among individuals who do not identify with the president's party.

Figure G2 illustrates the results of a multinomial probit model regressing vote choice on treatment condition, again without control variables (results including controls are substantively similar).

Figure G2. Change in Likelihood of Reported Voting Behavior by Treatment Condition



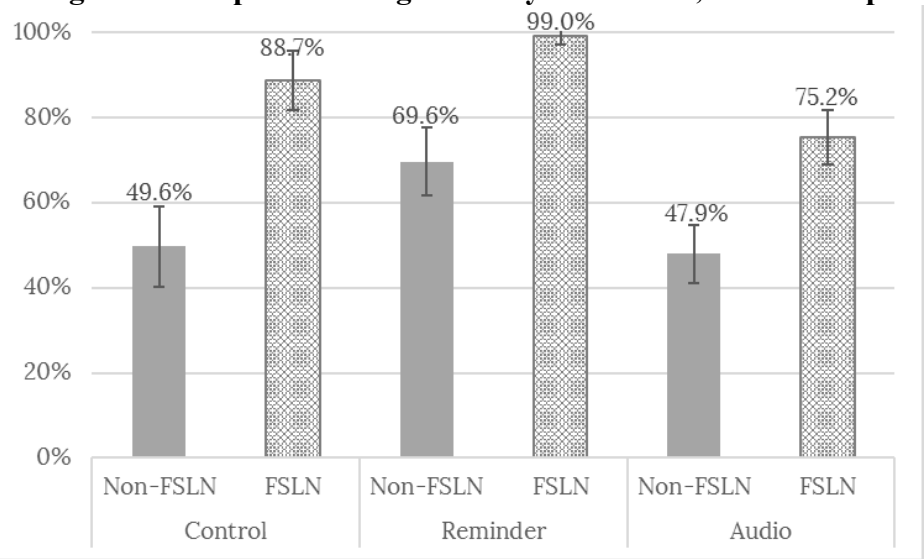
N=856. Dots represent the change in the predicted probability that a respondent will provide a given response when asked the vote question compared to the control condition. Vertical whiskers signify 95% confidence intervals around predicted probability. Probabilities were calculated using results from a multinomial probit estimated without controls.

Item non-response is significantly lower among respondents in both the Reminder (dark gray) and Audio (light gray) conditions, when compared to the Control group (the excluded category here). As above, respondents in the Reminder condition are 7.2 percentage points less likely to refuse to respond, and those in the Audio group are 10.9 percentage points less likely to refuse to answer the vote choice question, compared to those in the Control group. Consistent with perspectives linking audio-assisted interviewing to increased reporting of non-sanctioned behavior, respondents in the Audio group are 12.7 percentage points more likely to report casting an invalid ballot than individuals in the Control group.

Unexpectedly, respondents in the Reminder condition are 17 percentage points more likely to report a vote for Ortega than those in the Control and Audio conditions. Further, those in the Reminder condition are significantly less likely to report voting for a minor opposition candidate than those in the Control condition (8 percentage points) and the Audio condition (10 percentage points).

These effects are even more stark when broken down by partisanship: Figure G3 shows that those who identify with the incumbent FSLN party, as well as those who do not, are significantly more likely to report a vote for the incumbent in the reminder treatment than they are even in the control condition. This result suggests that the reminder backfired, especially for FSLN partisans, and that support for Ortega was probably over reported in non-audio treatments. Further, FSLN partisans report lower levels of support for Ortega, the election winner and their co-partisan, in the audio treatment conditions. Unfortunately, because the Nicaraguan government does not report complete disaggregated election results, it was not possible to assess which of these groups was being more truthful.

Figure G3. Reported Ortega Vote by Treatment, Partisanship



N=856. Predicted probabilities calculated following a multinomial logit model including non-response as an alternative. 95% confidence intervals shown.

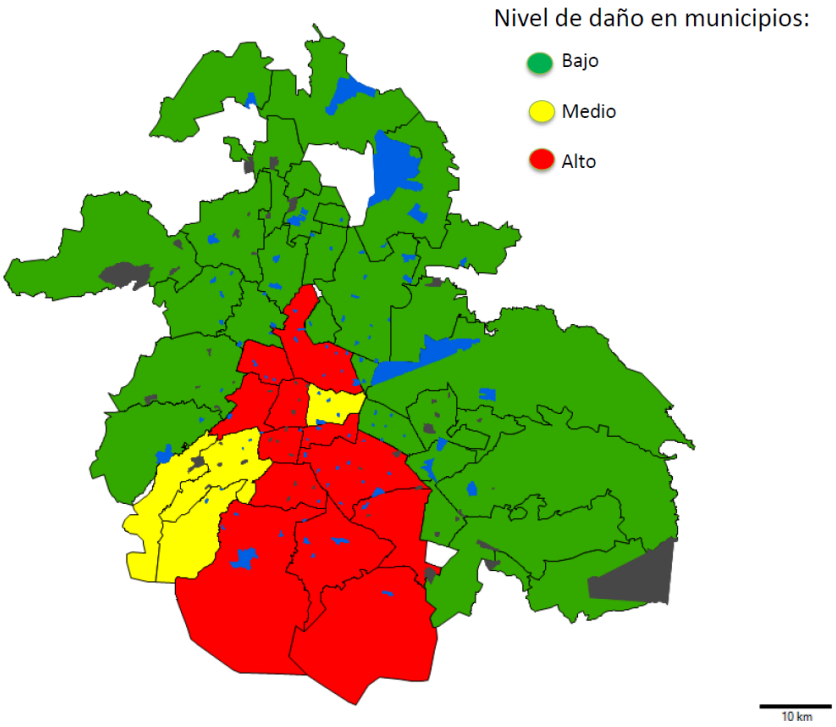
H. Sample design information and interviewer protocol

Sample Design México Post-Earthquake Study

Sample Design Post-Earthquake Study

On September 19, 2017, a magnitude 7.1 earthquake struck central Mexico and caused significant damage to the south parts of Mexico City and peripheral areas. As a result of the earthquake, more than 200 people were killed in Mexico City, thousands were injured, and numerous buildings and services collapsed. The map in Figure 1 shows the area affected by the earthquake.

Figure 1: Area affected by the earthquake



At the time of the September 19 earthquake, LAPOP together with the Inter-American Development Bank (IDB) was conducting a public opinion survey in the Federal District of Mexico and its Metropolitan Area.

The original sample design for the study comprised a total of 900 interviews of Mexican citizens of voting age who are residents of municipalities of the Federal District and its Metropolitan Area. At the time of the earthquake, almost two-thirds of the interviews (that is, 582 interviews of the 900) had been completed, and fieldwork as completed in 20 of the 45 municipalities or primary sampling units selected, while in 16 municipalities fieldwork had not commenced. Table 1 details the state of fieldwork at the time of the occurrence of the earthquake.

Table 1: Data collection status at the moment of the earthquake

Status	Number of Municipalities	Number of Interviews	Sample Size
Completed	20	540	540
Incomplete	9	42	252
Not started	16	0	192
Number of Municipalities (UPMs)	45	582	900

In order to obtain pre- and post-earthquake data from as homogeneous a sample as possible (in terms of their location within the geographical area, the sex and age of the interviewees and other exogenous factors to the treatment), so as to then evaluate the effects of this natural disaster on the public opinion, the previous sample was reproduced, but doubling the number of interviews in the areas (municipalities and clusters) where interviews had been conducted and completed at the time of the earthquake. The post-quake sample design includes the 20 municipalities or UPMs completed (540 interviews, doubled) and the 9 municipalities or UPMs partially interviewed at the time of the earthquake (42 interviews, doubled) plus the 318 interviews that were not completed in the pre-quake study. Table 2 describes the sample size of the pre- and post-earthquake studies

Table 2: Sample size in the Post-earthquake study

	Number of Municipalities	Number of Interviews Post-earthquake	Sample Size Post-earthquake
Completed	20	540	1082
Incomplete	9	42	84
Not started	16	0	318
Number of Municipalities (UPMs)	45	582	1,482

The idea behind doubling the size of the pre-earthquake interviews is to obtain additional margin in the probability of matching to the pre-earthquake interviewed sample. The new post-earthquake sample replicates the contours of the original sample, including with respect to the distribution of gender and age of the population of each municipality or UPM.
In the original study, the sample corresponds to a stratified, multi-stage probabilistic sampling by conglomerates. The design of the pre-earthquake study sample is described in the following section.

Sample design Pre-earthquake study

II.1 Universe, population and unit of observation

Universe: The universe is adults of voting age in the Metropolitan Area of Mexico City. The universe is comprised of adults (18 years and older) who live in the 76 municipalities of 3 states

according to the National Population Council and the 2010 Population and Housing Census of the National Institute of Geography and Information Statistics
Population: the survey is designed to collect information from a representative sample of the adult population of voting age in the Metropolitan Area of Mexico City; the population consists of permanent residents of area households and Mexican citizens.
Unit of observation: Individual respondent.

II.2 Sampling Frame

The sample is drawn using Electoral Districts (Secciones) of the National Electoral Institute (INE) as a reference frame of reference with updated data as of March 2016. In the specific case of this investigation, only the electoral districts (secciones) of the municipalities that make up the Metropolitan Area of Mexico City are considered.
The electoral districts (secciones) are considered as the Primary Sampling Units (UPM). For the territorial domain of this sample there are 8565 electoral sections. Table 1 shows the distribution of the population aged 18 and over according to municipalities and regions

Figure 3: Area of Study

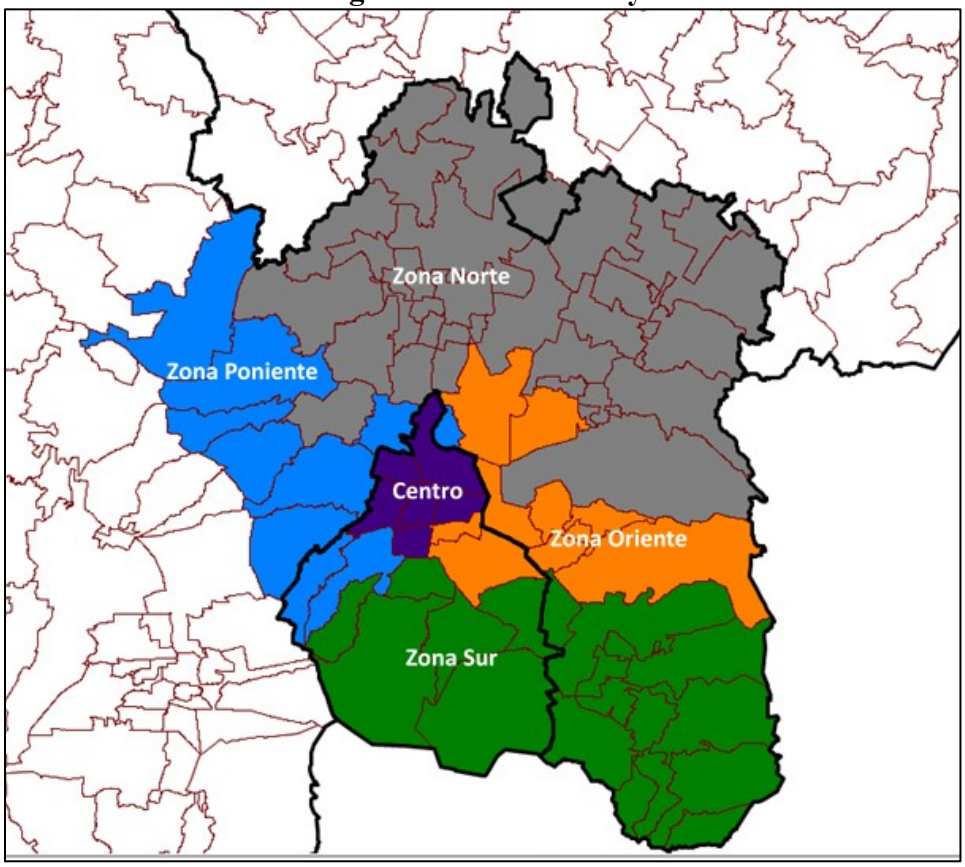


Table 3: Distribution by State, municipality and region of the study

State	Region	Municipality	Population	%Pob	Reg
Ciudad de México	Centro	Benito Juarez	385,439	1.9	1
Ciudad de México	Centro	Cuauhtemoc	531,831	2.6	1
Ciudad de México	Centro	Miguel Hidalgo	372,889	1.9	1
Ciudad de México	Centro	Azcapotzalco	414,711	2.1	1
Ciudad de México	Centro	Gustavo A. Madero	1,185,786	5.9	1
Ciudad de México	Centro	Venustiano Carranza	430,978	2.1	1
Ciudad de México	Sur	La Magdalena Contreras	239,086	1.2	2
Ciudad de México	Sur	Milpa Alta	130,582	0.6	2
Ciudad de México	Sur	Tlahuac	360,265	1.8	2
Ciudad de México	Sur	Coyoacan	620,416	3.1	2
Ciudad de México	Sur	Tlalpan	650,567	3.2	2
Ciudad de México	Sur	Xochimilco	415,007	2.1	2
Estado de México	Sur	Ayapango	8,864	0.0	2
Estado de México	Sur	Chalco	310,130	1.5	2
Estado de México	Sur	Cocotitlan	12,142	0.1	2
Estado de México	Sur	Ecatzingo	9,369	0.0	2
Estado de México	Sur	Juchitepec	23,479	0.1	2
Estado de México	Sur	Ozumba	27,207	0.1	2
Estado de México	Sur	Temamatla	11,206	0.1	2
Estado de México	Sur	Tenango del Aire	10,578	0.1	2
Estado de México	Sur	Tepetlixpa	18,327	0.1	2
Estado de México	Sur	Tlalmanalco	46,130	0.2	2
Estado de México	Sur	Amecameca	48,421	0.2	2
Estado de México	Sur	Valle de Chalco Solidaridad	357,645	1.8	2
Estado de México	Sur	Atlautla	27,663	0.1	2
Estado de México	Poniente	Cuajimalpa de Morelos	186,391	0.9	3
Estado de México	Poniente	alvaro Obregon	727,034	3.6	3
Estado de México	Poniente	Huixquilucan	242,167	1.2	3

State	Region	Municipality	Population	%Pob	Reg
Estado de México	Poniente	Isidro Fabela	10,308	0.1	3
Estado de México	Poniente	Jilotzingo	17,970	0.1	3
Estado de México	Poniente	Naucalpan de Juarez	833,779	4.1	3
Estado de México	Poniente	Nicolas Romero	366,602	1.8	3
Estado de México	Poniente	Tlalnepantla de Baz	664,225	3.3	3
Estado de México	Poniente	Villa del Carbon	44,881	0.2	3
Estado de México	Norte	Acolman	136,558	0.7	4
Estado de México	Norte	Apaxco	27,521	0.1	4
Estado de México	Norte	Atizapan de Zaragoza	489,937	2.4	4
Estado de México	Norte	Axapusco	25,559	0.1	4
Estado de México	Norte	Chiautla	26,191	0.1	4
Estado de México	Norte	Chiconcuac	22,819	0.1	4
Estado de México	Norte	Coacalco de Berriozabal	278,064	1.4	4
Estado de México	Norte	Coyotepec	39,030	0.2	4
Estado de México	Norte	Cuautitlan	140,059	0.7	4
Estado de México	Norte	Cuautitlan Izcalli	511,675	2.5	4
Estado de México	Norte	Huehuetoca	100,023	0.5	4
Estado de México	Norte	Hueypoxtla	39,864	0.2	4
Estado de México	Norte	Jaltenco	26,328	0.1	4
Estado de México	Norte	Melchor Ocampo	50,240	0.2	4
Estado de México	Norte	Nextlalpan	34,347	0.2	4
Estado de México	Norte	Nopaltepec	8,895	0.0	4
Estado de México	Norte	Otumba	34,232	0.2	4
Estado de México	Norte	Papalotla	4,147	0.0	4
Estado de México	Norte	San Martin de las Piramides	24,815	0.1	4
Estado de México	Norte	Tecamac	364,579	1.8	4
Estado de México	Norte	Temascalapa	35,987	0.2	4
Estado de México	Norte	Teoloyucan	63,115	0.3	4
Estado de México	Norte	Teotihuacan	53,010	0.3	4
Estado de México	Norte	Tepetlaoxtoc	27,944	0.1	4
Estado de México	Norte	Tepotzotlan	88,559	0.4	4
Estado de México	Norte	Tequixquiac	33,907	0.2	4
Estado de México	Norte	Tezoyuca	35,199	0.2	4
Estado de México	Norte	Texcoco	235,151	1.2	4
Estado de México	Norte	Tonanitla	10,216	0.1	4
Estado de México	Norte	Tultepec	91,808	0.5	4
Estado de México	Norte	Tultitlan	524,074	2.6	4

State	Region	Municipality	Population	%Pob	Reg
Estado de México	Norte	Zumpango	159,647	0.8	4
Hidalgo	Norte	Tizayuca	97,461	0.5	4
Ciudad de México	Oriente	Iztacalco	384,326	1.9	5
Ciudad de México	Oriente	Iztapalapa	1,815,786	9.0	5
Estado de México	Oriente	Atenco	56,243	0.3	5
Estado de México	Oriente	Chicoloapan	175,053	0.9	5
Estado de México	Oriente	Chimalhuacan	614,453	3.1	5
Estado de México	Oriente	Ecatepec de Morelos	1,656,107	8.2	5
Estado de México	Oriente	Ixtapaluca	467,361	2.3	5
Estado de México	Oriente	La Paz	253,845	1.3	5
Estado de México	Oriente	Nezahualcoyotl	1,110,565	5.5	5
		TOTAL	20,116,775	100.0	

II.3 Sampling method

The sampling method is designed to obtain a representative sample at the level of the Metropolitan Zone of Mexico City. For this, the sampling method corresponds to a stratified, multi-stage probability sampling by conglomerates. The first level of stratification consists in dividing into five strata the Metropolitan Zone of Mexico City into: ZMCM Norte, ZMCM Centro, ZMCM Este, ZMCM Sur y ZMCM Oeste, as indicated in the Table and Figure 3.
Stratified sampling ensures greater reliability in our sample by reducing the variance of the estimates, and, more importantly, ensures the inclusion of municipalities throughout the metropolitan area of Mexico City proportional to the size of the population of each. Stratification improves the quality of the estimates, with the only condition that the final sampling unit belongs to a single stratum and that the union of the strata conforms to the total population. Stratification also allows one to ensure dispersion of the sample. Table 4 indicates all the levels of stratification and sub-stratification of the sample:

Table 4: Probabilistic sampling, estratification

	ZMCM Norte, ZMCM Centro, ZMCM Este, ZMCM Sur y ZMCM Oeste
Primary Sampling Unit (PSU)	Electoral Sections
Secondary Sampling Unit (SSU)	Blocks
Tertiary Unit (TU)	Household
Final Unit	Respondent

First stage: stratification of the sample into 5 strata, according to the size of their population:

Region	Sampling points	Interviews	Percentage of Sampling
Centro	28	168	18.7%

Sur	25	150	16.7%
Poniente	23	138	15.3%
Norte	27	162	18.0%
Oriente	47	282	31.3%
TOTAL	**150**	**900**	**100**

The second stage, which corresponds to the selection of the Primary Sampling Units (PSUs), consists in the selection of Electoral Districts (Secciones) within each of the strata defined above with probability proportional to the adult population of voting age in the country.

The third stage in the sample design consists of the selection of clusters within each UPM using a systematic selection with probability proportion to the size of the population (PPS). We aim to conduct a maximum of 6 interviews in each of these Secondary Sampling Units (SSUs).

In the fourth stage, households are selected, which was done by counting the households in the selected SSUs and a systematic jump for the subsequent units.

Finally, in the fifth stage of the sample design, a frequency matching approach by sex and household is used to select a single interviewee in each household. The objective is to ensure that the distribution of individuals by sex and age in the survey corresponds to the population parameter that is used as a framework for the design of the sample. A completely random selection within the home would have required multiple attempts, drastically raising the costs but without any guarantee that at the end of multiple attempts it will have a correct balance of gender and age.

II.4 Sample Selection

First Stage: Sample stratification

The sample is stratified into 5 strata according to the size of its population (Table 5).

Table 5: Distribution of interviews by strata

Region	Sampling points	Interviews
Centro	28	168
Sur	25	150
Poniente	23	138
Norte	27	162
Oriente	47	282
TOTAL	**150**	**900**

Second Stage: Primary Sampling Units (PSU):

In the *s*econd stage, the Primary Sampling Units (PSU) are selected within each of the 5 strata. The Primary Sampling Unit in this study the Electoral District (Sección). As shown below, LAPOP completed the selection of Districts (Secciones) in the 5 strata that make up this sample.

REGION	STATE	SECCION	MUNICIPIO
CENTRO	CIUDAD DE MÉXICO	44	AZCAPOTZALCO
CENTRO	CIUDAD DE MÉXICO	133	AZCAPOTZALCO
CENTRO	CIUDAD DE MÉXICO	230	AZCAPOTZALCO
CENTRO	CIUDAD DE MÉXICO	329	AZCAPOTZALCO
CENTRO	CIUDAD DE MÉXICO	871	GUSTAVO A. MADERO
CENTRO	CIUDAD DE MÉXICO	941	GUSTAVO A. MADERO
CENTRO	CIUDAD DE MÉXICO	1037	GUSTAVO A. MADERO
CENTRO	CIUDAD DE MÉXICO	1151	GUSTAVO A. MADERO
CENTRO	CIUDAD DE MÉXICO	1243	GUSTAVO A. MADERO
CENTRO	CIUDAD DE MÉXICO	1336	GUSTAVO A. MADERO
CENTRO	CIUDAD DE MÉXICO	1424	GUSTAVO A. MADERO
CENTRO	CIUDAD DE MÉXICO	1518	GUSTAVO A. MADERO
CENTRO	CIUDAD DE MÉXICO	1596	GUSTAVO A. MADERO
CENTRO	CIUDAD DE MÉXICO	4273	BENITO JUAREZ
CENTRO	CIUDAD DE MÉXICO	4353	BENITO JUAREZ
CENTRO	CIUDAD DE MÉXICO	4425	BENITO JUAREZ
CENTRO	CIUDAD DE MÉXICO	4500	BENITO JUAREZ
CENTRO	CIUDAD DE MÉXICO	4590	CUAUHTEMOC
CENTRO	CIUDAD DE MÉXICO	4667	CUAUHTEMOC
CENTRO	CIUDAD DE MÉXICO	4761	CUAUHTEMOC
CENTRO	CIUDAD DE MÉXICO	4842	CUAUHTEMOC
CENTRO	CIUDAD DE MÉXICO	4941	MIGUEL HIDALGO

REGION	STATE	SECCION	MUNICIPIO
CENTRO	CIUDAD DE MÉXICO	5036	MIGUEL HIDALGO
CENTRO	CIUDAD DE MÉXICO	5121	MIGUEL HIDALGO
CENTRO	CIUDAD DE MÉXICO	5202	VENUSTIANO CARRANZA
CENTRO	CIUDAD DE MÉXICO	5290	VENUSTIANO CARRANZA
CENTRO	CIUDAD DE MÉXICO	5381	VENUSTIANO CARRANZA
CENTRO	CIUDAD DE MÉXICO	5464	VENUSTIANO CARRANZA
SUR	CIUDAD DE MÉXICO	369	COYOACAN
SUR	CIUDAD DE MÉXICO	433	COYOACAN
SUR	CIUDAD DE MÉXICO	509	COYOACAN
SUR	CIUDAD DE MÉXICO	594	COYOACAN
SUR	CIUDAD DE MÉXICO	668	COYOACAN
SUR	CIUDAD DE MÉXICO	2985	LA MAGDALENA CONTRERAS
SUR	CIUDAD DE MÉXICO	3069	LA MAGDALENA CONTRERAS
SUR	CIUDAD DE MÉXICO	3131	MILPA ALTA
SUR	CIUDAD DE MÉXICO	3627	TLAHUAC
SUR	CIUDAD DE MÉXICO	3677	TLAHUAC
SUR	CIUDAD DE MÉXICO	3728	TLAHUAC
SUR	CIUDAD DE MÉXICO	3804	TLALPAN
SUR	CIUDAD DE MÉXICO	3876	TLALPAN
SUR	CIUDAD DE MÉXICO	3945	TLALPAN
SUR	CIUDAD DE MÉXICO	3987	TLALPAN
SUR	CIUDAD DE MÉXICO	4073	TLALPAN

REGION	STATE	SECCION	MUNICIPIO
SUR	CIUDAD DE MÉXICO	4144	XOCHIMILCO
SUR	CIUDAD DE MÉXICO	4201	XOCHIMILCO
SUR	CIUDAD DE MÉXICO	4253	XOCHIMILCO
SUR	ESTADO DE MÉXICO	651	COCOTITLAN
SUR	ESTADO DE MÉXICO	963	VALLE DE CHALCO SOLIDARIDAD
SUR	ESTADO DE MÉXICO	1029	VALLE DE CHALCO SOLIDARIDAD
SUR	ESTADO DE MÉXICO	1042	CHALCO
SUR	ESTADO DE MÉXICO	1078	CHALCO
SUR	ESTADO DE MÉXICO	4751	TLALMANALCO
OESTE	CIUDAD DE MÉXICO	774	CUAJIMALPA DE MORELOS
OESTE	CIUDAD DE MÉXICO	3173	ALVARO OBREGON
OESTE	CIUDAD DE MÉXICO	3260	ALVARO OBREGON
OESTE	CIUDAD DE MÉXICO	3326	ALVARO OBREGON
OESTE	CIUDAD DE MÉXICO	3397	ALVARO OBREGON
OESTE	CIUDAD DE MÉXICO	3500	ALVARO OBREGON
OESTE	CIUDAD DE MÉXICO	3581	ALVARO OBREGON
OESTE	ESTADO DE MÉXICO	2018	HUIXQUILUCAN
OESTE	ESTADO DE MÉXICO	2064	HUIXQUILUCAN
OESTE	ESTADO DE MÉXICO	2635	NAUCALPAN DE JUAREZ
OESTE	ESTADO DE MÉXICO	2691	NAUCALPAN DE JUAREZ
OESTE	ESTADO DE MÉXICO	2779	NAUCALPAN DE JUAREZ
OESTE	ESTADO DE MÉXICO	2877	NAUCALPAN DE JUAREZ

REGION	STATE	SECCION	MUNICIPIO
OESTE	ESTADO DE MÉXICO	2936	NAUCALPAN DE JUAREZ
OESTE	ESTADO DE MÉXICO	3000	NAUCALPAN DE JUAREZ
OESTE	ESTADO DE MÉXICO	3734	NICOLAS ROMERO
OESTE	ESTADO DE MÉXICO	3775	NICOLAS ROMERO
OESTE	ESTADO DE MÉXICO	3816	NICOLAS ROMERO
OESTE	ESTADO DE MÉXICO	4802	TLALNEPANTLA DE BAZ
OESTE	ESTADO DE MÉXICO	4868	TLALNEPANTLA DE BAZ
OESTE	ESTADO DE MÉXICO	4946	TLALNEPANTLA DE BAZ
OESTE	ESTADO DE MÉXICO	5026	TLALNEPANTLA DE BAZ
OESTE	ESTADO DE MÉXICO	5112	TLALNEPANTLA DE BAZ
NORTE	ESTADO DE MÉXICO	50	ACOLMAN
NORTE	ESTADO DE MÉXICO	265	ATIZAPAN DE ZARAGOZA
NORTE	ESTADO DE MÉXICO	302	ATIZAPAN DE ZARAGOZA
NORTE	ESTADO DE MÉXICO	351	ATIZAPAN DE ZARAGOZA
NORTE	ESTADO DE MÉXICO	408	ATIZAPAN DE ZARAGOZA
NORTE	ESTADO DE MÉXICO	556	COACALCO DE BERRIOZABAL
NORTE	ESTADO DE MÉXICO	607	COACALCO DE BERRIOZABAL
NORTE	ESTADO DE MÉXICO	675	CUAUTITLAN
NORTE	ESTADO DE MÉXICO	725	CUAUTITLAN IZCALLI
NORTE	ESTADO DE MÉXICO	801	CUAUTITLAN IZCALLI
NORTE	ESTADO DE MÉXICO	864	CUAUTITLAN IZCALLI
NORTE	ESTADO DE MÉXICO	1970	HUEHUETOCA

REGION	STATE	SECCION	MUNICIPIO
NORTE	ESTADO DE MÉXICO	2459	MELCHOR OCAMPO
NORTE	ESTADO DE MÉXICO	4193	TECAMAC
NORTE	ESTADO DE MÉXICO	4228	TECAMAC
NORTE	ESTADO DE MÉXICO	4507	TEOLOYUCAN
NORTE	ESTADO DE MÉXICO	4567	TEPOTZOTLAN
NORTE	ESTADO DE MÉXICO	4629	TEXCOCO
NORTE	ESTADO DE MÉXICO	4676	TEXCOCO
NORTE	ESTADO DE MÉXICO	5465	TULTEPEC
NORTE	ESTADO DE MÉXICO	5503	TULTITLAN
NORTE	ESTADO DE MÉXICO	5557	TULTITLAN
NORTE	ESTADO DE MÉXICO	5606	TULTITLAN
NORTE	ESTADO DE MÉXICO	5882	ZUMPANGO
NORTE	ESTADO DE MÉXICO	5921	ZUMPANGO
NORTE	ESTADO DE MÉXICO	6288	TECAMAC
NORTE	ESTADO DE MÉXICO	6464	CUAUTITLAN IZCALLI
ORIENTE	CIUDAD DE MÉXICO	1688	IZTACALCO
ORIENTE	CIUDAD DE MÉXICO	1777	IZTACALCO
ORIENTE	CIUDAD DE MÉXICO	1867	IZTACALCO
ORIENTE	CIUDAD DE MÉXICO	1955	IZTACALCO
ORIENTE	CIUDAD DE MÉXICO	2042	IZTAPALAPA
ORIENTE	CIUDAD DE MÉXICO	2125	IZTAPALAPA
ORIENTE	CIUDAD DE MÉXICO	2200	IZTAPALAPA

REGION	STATE	SECCION	MUNICIPIO
ORIENTE	CIUDAD DE MÉXICO	2274	IZTAPALAPA
ORIENTE	CIUDAD DE MÉXICO	2359	IZTAPALAPA
ORIENTE	CIUDAD DE MÉXICO	2445	IZTAPALAPA
ORIENTE	CIUDAD DE MÉXICO	2541	IZTAPALAPA
ORIENTE	CIUDAD DE MÉXICO	2612	IZTAPALAPA
ORIENTE	CIUDAD DE MÉXICO	2677	IZTAPALAPA
ORIENTE	CIUDAD DE MÉXICO	2742	IZTAPALAPA
ORIENTE	CIUDAD DE MÉXICO	2802	IZTAPALAPA
ORIENTE	CIUDAD DE MÉXICO	2864	IZTAPALAPA
ORIENTE	CIUDAD DE MÉXICO	2921	IZTAPALAPA
ORIENTE	ESTADO DE MÉXICO	240	ATENCO
ORIENTE	ESTADO DE MÉXICO	1176	CHIMALHUACAN
ORIENTE	ESTADO DE MÉXICO	1215	CHIMALHUACAN
ORIENTE	ESTADO DE MÉXICO	1267	CHIMALHUACAN
ORIENTE	ESTADO DE MÉXICO	1342	ECATEPEC DE MORELOS
ORIENTE	ESTADO DE MÉXICO	1382	ECATEPEC DE MORELOS
ORIENTE	ESTADO DE MÉXICO	1440	ECATEPEC DE MORELOS
ORIENTE	ESTADO DE MÉXICO	1513	ECATEPEC DE MORELOS
ORIENTE	ESTADO DE MÉXICO	1572	ECATEPEC DE MORELOS
ORIENTE	ESTADO DE MÉXICO	1660	ECATEPEC DE MORELOS
ORIENTE	ESTADO DE MÉXICO	1733	ECATEPEC DE MORELOS
ORIENTE	ESTADO DE MÉXICO	1778	ECATEPEC DE MORELOS

REGION	STATE	SECCION	MUNICIPIO
ORIENTE	ESTADO DE MÉXICO	1829	ECATEPEC DE MORELOS
ORIENTE	ESTADO DE MÉXICO	1889	ECATEPEC DE MORELOS
ORIENTE	ESTADO DE MÉXICO	2078	IXTAPALUCA
ORIENTE	ESTADO DE MÉXICO	2143	IXTAPALUCA
ORIENTE	ESTADO DE MÉXICO	3047	NEZAHUALCOYOTL
ORIENTE	ESTADO DE MÉXICO	3129	NEZAHUALCOYOTL
ORIENTE	ESTADO DE MÉXICO	3218	NEZAHUALCOYOTL
ORIENTE	ESTADO DE MÉXICO	3293	NEZAHUALCOYOTL
ORIENTE	ESTADO DE MÉXICO	3385	NEZAHUALCOYOTL
ORIENTE	ESTADO DE MÉXICO	3476	NEZAHUALCOYOTL
ORIENTE	ESTADO DE MÉXICO	3558	NEZAHUALCOYOTL
ORIENTE	ESTADO DE MÉXICO	3644	NEZAHUALCOYOTL
ORIENTE	ESTADO DE MÉXICO	3945	LA PAZ
ORIENTE	ESTADO DE MÉXICO	3984	LA PAZ
ORIENTE	ESTADO DE MÉXICO	5974	CHIMALHUACAN
ORIENTE	ESTADO DE MÉXICO	6006	ECATEPEC DE MORELOS
ORIENTE	ESTADO DE MÉXICO	6097	IXTAPALUCA
ORIENTE	ESTADO DE MÉXICO	6372	CHICOLOAPAN

Third Stage: selection of blocks

In the third stage of the sample selection process, housing blocks are selected in each PSU, with a systematic selection with probability proportional to the size of each element. That is, housing blocks are selected according to a Probability Proportional to the Size of the Population (PPT) in a systematic way with a random starting point within each PSU.

Fourth Stage: selection of households

This stage of the selection begins once the interviewers locate the starting point of the housing block (point north-east of the block or block, and walking in the clockwise direction). Each interviewer will choose a number of households in a systematic way. Specifically, interviews should be conducted every three homes. In other words, each time an interview is completed, the following interview cannot be conducted in the next two homes.

In case of rejection, unoccupied dwelling or absence of people, the interviewer will select the adjacent dwelling. In cases where an interviewer reaches the end of a block of houses without having completed the quota of interviews, he or she can proceed to the next block, following the same routine as in the previous block.

Fifth Stage: selection of respondents

Finally, a person to be interviewed in each household will be selected. The frequency match for each age group and sex was estimated based on the distribution of the population registered in the Electoral Districts (Secciones) and with formats A, B and C to get as close as possible to the reference parameter, the quota forms were interspersed between the points. Initially, one third of the sample was controlled in the A format, another third in the B format, and the final third in the C format. The interviewee must be a permanent member of the household, not a domestic or visitor job. If there are more than two people in the same age group and sex in the home, the questionnaire should be applied to the person with the closest birthday.

Table 6: Gender and Age Interviews Distribution

Group A				
Gender/Age	**18- 29**	**30- 50**	**50 and over**	**Total**
Male	1	2	0	3
Female	1	1	1	3
Total	2	3	1	6
Group B				
Gender/Age	**18- 29**	**30- 50**	**50 and over**	**Total**
Male	1	1	1	3
Female	1	2	0	3
Total	2	3	1	6
Group C				
Gender/Age	**18- 29**	**30- 50**	**50 and over**	**Total**
Male	1	1	1	3
Female	1	1	1	3
Total	2	2	2	6

II.5 Level of confidence and margin of error

The level of confidence anticipated for the national sample is 95%, with a margin of error of 3.3 percent, assuming a 50/50 ratio in the dichotomous variables (in any other proportion, the sampling error is lower) and a Simple Random Sample (MAS). Given that the sample is stratified and by conglomerates (Kish 1995), the complex sample design has to be taken into account in order to accurately estimate the precision of the sample. It is not possible to determine the sampling error *a priori.*