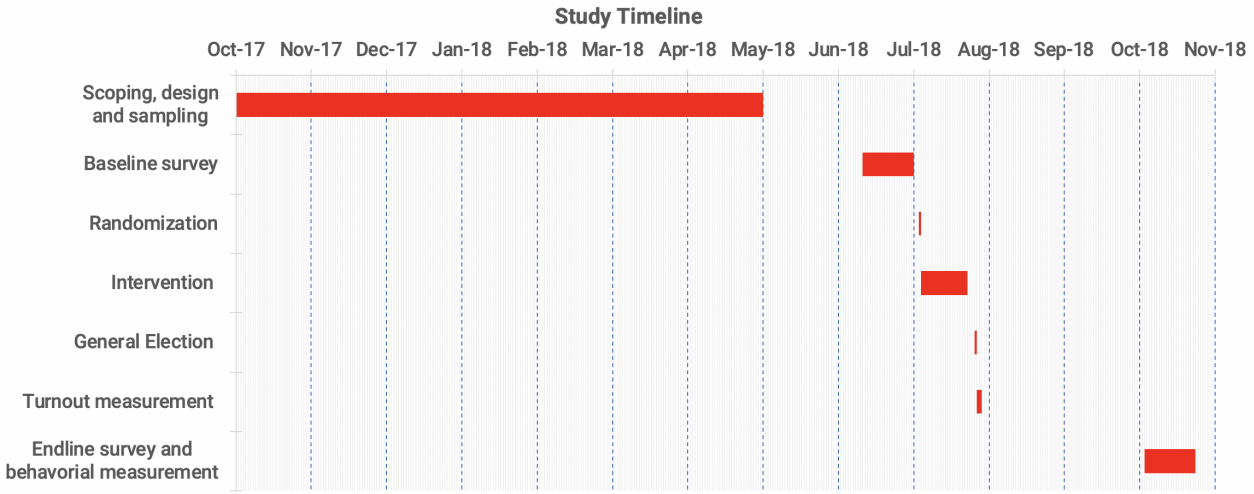
Appendix to
Canvassing the Gatekeepers: A Field Experiment to Increase Women Voters' Turnout in Pakistan

Contents

A Study Design and Implementation

A.1 Study Timeline



A.2 Sampling Strategy

Our sample of study households is drawn from 94 geographically contiguous Union Councils (local administrative unit) in the northern part of the city of Lahore. Each Union Council has 6 wards within it. The ward is the lowest administrative and political unit and serves as the unit of randomization. To draw a sample of 500 wards, we include all 6 wards from a random subset of 30 Union Councils, and then randomly select 5 out of 6 wards for inclusion in the study from the remaining 64 Union Councils. To obtain the sample of 2,500 households, we select 5 households in each of the 500 sample wards.

To sample households within a ward, we drop a location pin at a random point within each ward. Two enumerators (one woman and one man) proceed to the pin location, and select the nearest household to the right for the first survey. After completing the first survey, they select four other households in the ward using the right hand rule which involves selecting the 7th household to the right of the last household surveyed. A household is excluded from the sample if the dwelling is locked/empty, if all members of the household are not registered to vote, if all members are registered to vote outside of Lahore, or if there is not at least 1 adult woman and 1 adult man with a CNIC (Computerized National Identity Card, which is required to vote) available and consenting to be surveyed. These conditions restrict the sample to households with individuals who could plausibly cast a vote (have a CNIC and are registered in Lahore) because our intervention is conducted after the preparation of electoral rolls, which means we cannot expect it to change voter registration status.

If a household is ineligible for inclusion for any of these reasons, the enumerators skip the dwelling and proceed to one immediately to the right of it. Within the household, respondents are selected by listing all N eligible (over the age of 18 and possessing a CNIC)

respondents of the same gender as the enumerator in order of age. After the listing is complete, a random number generator programmed in the survey tablet generates a number n, and the enumerator asks to speak with the *nth* listed eligible individual. The enumerator then provides relevant information about the study to this selected individual using an information script, and seeks oral consent to conduct a survey with them (see Appendix A.6 for details).

Households that participate in the study received compensation of Rs.100 (approximately $0.8 at the average exchange rate in 2018) after completion of the baseline survey, where completion means that the randomly selected female and male individuals have completed the survey, and another Rs.100 upon completion of the endline survey. The compensation was delivered via mobile money transfer if a member of the household had chosen to provide a cellphone number, or in the form of a physical mobile money scratch card to households where participants had chosen not to share a cellphone number. The minimum wage in Pakistan is Rs.15,000 per month, less than Rs.100 per hour (estimating twenty 8-hour workdays in a month). Participant households, where two respondents completed a half-hour long survey in each phase, were therefore compensated at a comparable hourly rate. This compensation was presented as a "token of appreciation" for the respondents' time.

A.3 Power Calculations

We calculate the minimum detectable effects (MDE) for our study design using the STATA "power twoproportions, cluster" command which estimates effect size for a test comparing two independent proportions in cluster randomized designs. We estimate MDE for two comparisons:

1. Main effect of either treatment arm (corresponds to the specification reported in Cols 2-3 in Table 5)
2. Comparison of any treatment condition to control (corresponds to the specification reported in Col 1 in Table 5)

For both power calculations the following common study parameters:

- α (significance level) = 0.05
- κ (power) = 0.80
- P_0 (control mean) = 0.59 [1]
- ρ (intra-cluster correlation)=0.07 [2]

For the first comparison we use the following cluster design parameters, and estimate an **MDE=0.067**

- k1 (number of clusters in group 1) = 250 [3]

[1] This is the control group mean of self reported turnout among women respondents in our baseline survey
[2] This is the ICC on the self reported turnout variable among women in our baseline survey data
[3] This pools the 125 control clusters and 125 clusters assigned to T1only or T2only

- k2 (number of clusters in group 2) = 250 [4]
- m1 (observations per cluster in group 1) = 4 [5]
- m1 (observations per cluster in group 2) = 4

For the second comparison we use the following cluster design parameters, and estimate an **MDE=0.09**

- k1 (number of clusters in control) = 125
- k2 (number of clusters in any treatment condition) = 125
- m1 (observations per cluster in control) = 5
- m1 (observations per cluster in any treatment condition) = 4

A.4 Step-by-Step Details of Intervention

Step 1: Approaching Household Members
A female (male) canvasser visits a treatment household unannounced and requests to speak with the female (male) individual from that household who was surveyed at baseline for a 20-minute conversation.[6] If the baseline respondent is unavailable, the canvasser inquires when they might be home and if able to secure a time for later in the same day, they moved on to other households in the same area and return to the household later. If unable to make contact with the baseline respondent after 3 attempts, they ask to speak with any adult household member of the same gender as the baseline respondent. The canvasser asks the baseline respondent (or other individual) to gather all adult household members of the same gender who are available for a 20-minute conversation.

Step 2: Canvasser Introduction
Canvasser explicitly states their CSO affiliation, that they are here to speak about women's participation in the upcoming election, and clarifies that their organization is non-partisan. The canvassers also show letters from the Election Commission of Pakistan stating that this is an approved activity by a non-partisan organization.

Step 3: Motivational Video
The canvasser uses a handheld tablet device to show a 5-minute long video to household members. The video follows the narrative of a young woman, facing issues of poor service delivery in her neighborhood. This woman decides to make her voice heard by contacting a

[4] This pools the 125 clusters assigned to T1only or T2only and the 125 clusters assigned to T1+T2 both

[5] There are 5 control observations in each control cluster, but 4 treatment observations + 1 within-treatment control observation in each treatment cluster. Since this group pools together control and treatment, we go with the more conservative number

[6] Although this intervention is delivered at the household level, canvassers are asked to prioritize inclusion of the baseline survey respondent in an effort to maximize the chances that when we re-survey this respondent at endline, we are speaking to a household member who was present at the time of the intervention.

political candidate, and casting her vote in the election. Her brother is shown in an enabling role: he encourages her to take action and also agrees to help the women in his family reach the women's polling station on his motorbike.

The video content is designed in the spirit of "edutainment" style interventions. The video depicts a commonly experienced neighborhood problem (sanitation and sewerage) at the outset to make the video instantly relatable to viewers. The video then emphasizes the potential instrumental advantage of achieving tangible change through holding politicians accountable on election-day. This choice is informed by observations in focus groups whereby participants frequently expressed dissatisfaction with politician performance as a reason for disengagement with politics.

The video also models supportive behavior by male family members whereby the male character expresses verbal support of his sister's political participation, encourages his mother to vote in the election, and also provides tangible support to do so by offering to take his female family members to the polling station on his motorbike.[7]

Step 4: Procedural Information
The canvasser shares procedural and practical information about the election and voting process through informational leaflets and a practical demonstration of how to cast a ballot. The leaflets (see the next two sections) describe how to find out the location of one's polling booth, the process of voting and associated rules as well as the role of elected officials at the national and provincial level and the symbols assigned to various parties. The canvasser goes through all the information provided in the leaflets in person, and leaves copies of the leaflets with the household members. Then the canvasser uses mock ballot papers, ballot boxes and a stamp to show the household members exactly how to mark the ballot, fold the paper and put it in the ballot box.

[7]Our baseline data show that women are less likely to be encouraged to vote by family members than are men, and focus groups suggested that motorbike was the most common way of transportation to the polls in Lahore, and this is borne out in our data where nearly 70% of those who report having voted say that they reached the polling station by motorbike.

A.5 Informational Materials used in Intervention

Figure A.1: Procedural Information Leaflet

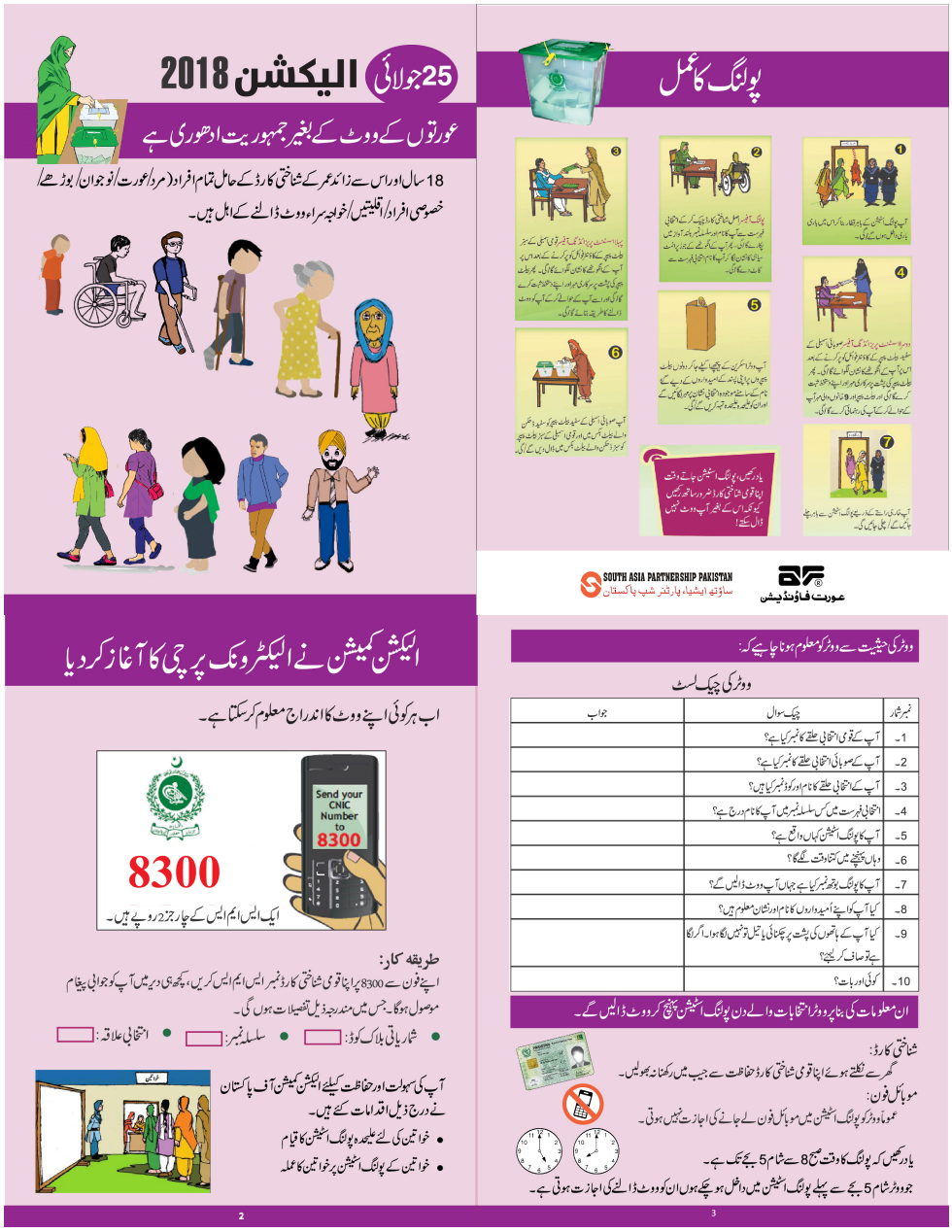


Figure A.2: Political Knowledge Leaflet

A.6 Ethical Considerations

Participant Information and Consent

Below is the translated information script used to obtain oral consent from study participants during data collection activities:

Hello, my name is []. I am here on behalf of researchers from [institution] and would like to invite you to participate in a survey. The reason why we are conducting this survey is to find out what people think about different political issues, what their service delivery priorities are and how decisions are made in their households. Your household has been selected through a randomization procedure. We would like to survey one male and one female member in each house. Only those males and females who are above the age of 18 and have CNIC's are eligible to participate in this survey

You are free to choose whether or not to participate in this survey. If you do choose to participate, I will require half an hour of your time. During the survey you can refuse to answer any questions that you do not wish answer, or ask me to end the interview at any point.

I also want to clarify that the information you provide us will only be used for research purposes. If you participate, we will retain your name, address, and phone number for a little while so that our firm can return again to ask you some more questions. We will carefully safeguard this information and store it securely and will not share it with anyone.

We cannot guarantee that you will benefit from this survey and research, but we will provide Rs. 100 in the form of mobile credit as a token of our appreciation if you decide to participate in this survey.

Do you have any questions? Would you like to participate? [Yes/No]

If you have any questions regarding the survey, after I leave, you can contact [local researcher phone number]. I will also leave a copy of this contact information with you after the survey.

Below is the translated introductory script used by canvassers conducting the intervention:

Hello, my name is [X] and I work with Aurat Foundation/South Asia Partnership Pakistan. Our organization works for women's political, social and economic progress in Pakistan. As you know the General Elections are coming up soon and I am here, on behalf of the Election Commission, to provide you with some information about the election, and discuss the importance of voting with you and your family. Would it be possible to speak with all the adult men (if male canvasser)/women (if female canvasser) who are present at home right now? I will take 20 minutes of your time.

While the information provided to study participants clearly discloses that they are being

asked to participate in a research study about politics, participants are not made aware that the survey team is assessing the impact of the intervention, or that the intervention is part of the study, or about their assignment to treatment or control status.

Electoral Intervention

We elaborate here on the specific ethical concerns associated with an intervention that is delivered around an election: the likelihood that such an intervention could affect the results of an election (Slough, 2020).
Our intervention is delivered across 7 electoral constituencies. The maximum number of treated households in any single constituency is 676.
The average number of adults in a household in urban Lahore is 3.45 (Pakistan Social and Living Standards Measurement Survey 2015-2016). If we take 4 as an upper bound, we would expect that a maximum of 2704 adult individuals would be treated in any one constituency in our sample.

In the previous national election in 2013, the lowest margin of victory in any constituency in Lahore was 7453 votes.

If we make an extreme assumption that none of the directly treated adults would have voted in absence of our intervention, we would need to see treatment effects amounting to a 275 percentage point increase in turnout to come close to the number of votes needed to swing an election. Given that most GOTV interventions achieve closer to 10 percentage point increases in turnout (Giné and Mansuri, 2018), we consider it highly unlikely that our intervention could plausibly change electoral outcomes.

B Outcome Measures

B.1 Behavioral Measure of Men's Supportive Behavior

Figure B.1: Stickers offered to Male Respondents at Endline

Note Sticker 1 (L) has the slogan "Strong Democracy, Strong Pakistan; Strong Democracy, Strong Pakistan"; Sticker 2 (R) has the slogan "Strong Democracy, Strong Pakistan; Without Women's Participation, Democracy is Incomplete"

C Effects on Index Measures and Index Components

C.1 All Indices

Table C.1: Results: Knowledge, Attitudes and Behaviors (ITT, Survey Measures)

	Panel A: Women's Responses					
	(1) **Political Knowledge**	(2) **Interest in Politics**	(3) **Self Efficacy**	(4) **Views on Men's Restrictions**	(5) **Election-Day Help by Men**	(6) **Political Discussion**
T1 only	-0.039	0.028	-0.020	0.048	-0.070	-0.004
	(0.073)	(0.055)	(0.078)	(0.069)	(0.073)	(0.033)
T2 only	-0.078	0.089	-0.002	-0.008	-0.032	0.036
	(0.082)	(0.055)	(0.083)	(0.074)	(0.079)	(0.035)
T1+T2	-0.101	0.039	-0.013	0.087	0.158**	0.059*
	(0.079)	(0.058)	(0.078)	(0.071)	(0.078)	(0.034)
Within T Ctrl	-0.047	0.008	0.052	0.037	0.090	0.030
	(0.070)	(0.048)	(0.073)	(0.062)	(0.065)	(0.032)
Constant	-1.168***	-0.394***	-0.423	0.414	0.204	0.617***
	(0.256)	(0.149)	(0.294)	(0.359)	(0.176)	(0.080)
UC FEs	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.184	0.244	0.134	0.368	0.236	0.141
# Observations	2433	2435	2431	2430	2381	2500
	Panel B: Men's Responses					
T1 only	-0.013	0.015	0.019	0.055	0.055	0.067*
	(0.038)	(0.070)	(0.065)	(0.080)	(0.071)	(0.035)
T2 only	-0.009	-0.010	-0.030	0.067	0.002	0.013
	(0.042)	(0.069)	(0.073)	(0.085)	(0.077)	(0.036)
T1+T2	-0.052	0.133*	0.063	-0.070	0.173**	0.077**
	(0.046)	(0.074)	(0.070)	(0.079)	(0.076)	(0.037)
Within T Ctrl	-0.019	0.047	-0.072	-0.033	0.078	0.013
	(0.040)	(0.064)	(0.065)	(0.068)	(0.069)	(0.033)
Constant	0.704***	0.563***	-0.229	0.025	0.725***	0.445***
	(0.095)	(0.124)	(0.149)	(0.447)	(0.118)	(0.125)
UC FEs	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.155	0.260	0.105	0.304	0.185	0.115
# Observations	2433	2434	2433	2433	2431	2499

Notes: All specifications show results using OLS estimation and employ block (Union Council) fixed effects. Standard errors in parentheses are clustered at the ward level. All outcomes are standardized indices, except for column (6). Column (6) is a binary indicator for whether women (men) stated that they discussed politics with a man (woman) in the household. For the remaining five columns, definitions of the variables composing the indices and results on each individual component are included in appendix tables Table C.4 to C.8.* p<0.10, ** p<0.05, *** p<0.01

C.2 Multiple Comparisons

Table C.2 shows results on the survey outcomes for the pooled sample of all respondents (men and women). To account for the fact that we are considering 18 different comparisons, six for each of the three treatment groups, we report adjusted test statistics in Table C.3 using several approaches.

The first row reports the unadjusted p-values from Table C.2. The second row reports the p-values using the Bonferroni correction: we multiply the unadjusted p-values by the total number of tests (18) and cap the maximum value at 1. Using this correction, we see that the p-value on the estimated effect of T1+T2 on election day help by men is below 0.10; however the p-value on the effect of T1+T2 on political discussion is above 0.10.

The third row reports adjusted p-values using the Benjamini-Hochberg procedure – to compute these the "raw" p-values are multiplied by m/i where m is the number of tests (18) and i is the rank of the p-value when p-values are sorted in ascending order (the smallest p-value has a rank of 1). If the adjusted p-value is smaller than the false discovery rate (i.e. the expected proportion of rejections that are type I errors or false rejections,FDR), the test is significant. Using this correction, the estimated effect of T1+T2 on election day help by men and political discussion is significant assuming a FDR of 10%. The fourth row reports sharpened False Discovery Rate (FDR) q-values following the approach in Anderson (2008). These are directly comparable to the raw p-values and we find that the sharpened q-values for the estimated effect of T1+T2 on election day help by men and political discussion are below 0.10.

Table C.2: Results: Knowledge, Attitudes and Behaviors (ITT, Survey Measures)

	(1) **Political Knowledge**	(2) **Interest in Politics**	(3) **Self Efficacy**	(4) **Views on Men's Restrictions**	(5) **Election-Day Help by Men**	(6) **Political Discussion**
T1 only	-0.027	0.021	-0.001	0.053	-0.009	0.031
	(0.042)	(0.049)	(0.050)	(0.052)	(0.053)	(0.026)
T2 only	-0.044	0.039	-0.017	0.031	-0.014	0.024
	(0.048)	(0.048)	(0.055)	(0.054)	(0.055)	(0.025)
T1+T2	-0.077	0.086*	0.025	0.010	0.165***	0.068***
	(0.047)	(0.050)	(0.052)	(0.054)	(0.057)	(0.026)
Within T Ctrl	-0.034	0.027	-0.010	0.004	0.084*	0.022
	(0.042)	(0.043)	(0.047)	(0.044)	(0.050)	(0.023)
Constant	-0.231*	0.084	-0.326***	0.218***	0.465***	0.531***
	(0.131)	(0.130)	(0.110)	(0.064)	(0.105)	(0.063)
UC FEs	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.074	0.103	0.069	0.189	0.120	0.077
# Observations	4866	4869	4864	4863	4812	4999

Notes: All specifications show results using OLS estimation and employ block (Union Council) fixed effects. Standard errors in parentheses are clustered at the ward level. All outcomes are standardized indices, except for column (6). Column (6) is a binary indicator for whether women (men) stated that they discussed politics with a man (woman) in the household. For the remaining five columns, definitions of the variables composing the indices and results on each individual component are included in appendix tables Table C.4 to C.8.* p<0.10, ** p<0.05, *** p<0.01

Table C.3: Adjusted test statistics for index outcomes using various approaches to multiple comparisons

	(1) **Political Knowledge**	(2) **Interest in Politics**	(3) **Self Efficacy**	(4) **Views on Restrictions**	(5) **Election Day Help by Men**	(6) **Political Discussion**
T1 only						
p-value	0.526	0.667	0.983	0.309	0.862	0.222
Bonferroni corrected p-value	1	1	1	1	1	1
Benjamini-Hochberg adjusted p-value	0.947	0.924	0.983	0.927	0.913	0.799
sharpened q-value	1	1	1	1	1	1
T2 only						
p-value	0.356	0.417	0.758	0.566	0.795	0.329
Bonferroni corrected p-value	1	1	1	1	1	1
Benjamini-Hochberg adjusted p-value	0.801	0.834	0.975	0.926	0.954	0.846
sharpened q-value	1	1	1	1	1	1
T1+T2						
p-value	0.104	0.085	0.626	0.854	0.004	0.010
Bonferroni corrected p-value	1	1	1	1	0.072	0.18
Benjamini-Hochberg adjusted p-value	0.468	0.51	0.939	0.961	0.072	0.09
sharpened q-value	0.713	0.713	1	1	0.078	0.093
N	4866	4869	4864	4863	4812	4999

C.3 Effects on Individual Index Components

Table C.4: Treatment Effects on Political Knowledge

	(1) ECP Phone	(2) Election Days	(3) Voter Signature	(4) PO Signature	(5) Party Slogans
Panel A: Women's Knowledge					
T1only	-0.036	-0.030	0.022	-0.076**	0.032
	(0.030)	(0.032)	(0.033)	(0.032)	(0.075)
T2only	-0.049	-0.016	0.024	-0.069*	-0.032
	(0.031)	(0.034)	(0.035)	(0.038)	(0.087)
T1+T2	-0.015	-0.028	0.034	-0.079**	-0.074
	(0.033)	(0.034)	(0.034)	(0.034)	(0.084)
Within T Control	-0.042	0.014	0.052*	-0.042	-0.059
	(0.028)	(0.030)	(0.031)	(0.032)	(0.073)
Constant	0.186***	0.337***	0.862***	0.090***	-0.948**
	(0.070)	(0.130)	(0.073)	(0.033)	(0.375)
UC FEs	Yes	Yes	Yes	Yes	Yes
R-Squared	0.126	0.248	0.097	0.155	0.230
# Observations	2423	2421	2408	2417	2428
Panel B: Men's Knowledge					
T1only	0.081**	-0.032**	-0.005	-0.075**	0.008
	(0.032)	(0.015)	(0.037)	(0.038)	(0.033)
T2only	0.094***	-0.018	-0.037	-0.062	0.007
	(0.034)	(0.015)	(0.038)	(0.040)	(0.037)
T1+T2	0.066*	-0.011	-0.047	-0.068*	-0.030
	(0.034)	(0.014)	(0.039)	(0.039)	(0.044)
Within T Control	0.027	-0.015	-0.002	-0.055	0.006
	(0.032)	(0.014)	(0.033)	(0.035)	(0.035)
Constant	0.710***	1.011***	0.567***	0.531***	0.606***
	(0.172)	(0.010)	(0.170)	(0.083)	(0.028)
UC FEs	Yes	Yes	Yes	Yes	Yes
R-Squared	0.204	0.084	0.264	0.220	0.147
# Observations	2427	2433	2413	2431	2433

Notes: All specifications show results using OLS estimation and employ strata (Union Council) fixed effects. Standard errors in parentheses are clustered at the ward level. All five outcomes used in this table are combined in a standardized index to form the outcome variable for Column (1) of Table C.1. Outcome for column (1) is an indicator for whether the respondent correctly repeated the Election Commission of Pakistan SMS short-code for checking one's voter registration. Column (2) is an indicator for whether the respondent correctly stated that elections for provincial and national assemblies take place on the same day (as opposed to different days). Column (3) is an indicator for whether the respondent correctly stated that a voter's signature is not required on the ballot paper. Column (4) is an indicator for whether the respondent correctly stated that a Presiding Officer's signature are required on the ballot paper. Column (5) is a standardized index comprising of four variables, each being an indicator for whether the respondent correctly linked a popular political slogan with a political party. * p<0.10, ** p<0.05, *** p<0.01

Table C.5: Treatment Effects on Interest in Politics

	Panel A: Women's Interest		
	(1) **Interest in Political TV**	(2) **Interest in Political Issues**	(3) **Interest in 2018 Election**
T1only	0.061	0.055	-0.047
	(0.056)	(0.055)	(0.070)
T2only	0.028	0.117**	0.083
	(0.054)	(0.055)	(0.074)
T1+T2	0.046	0.094	-0.037
	(0.059)	(0.058)	(0.075)
Within T Control	0.012	0.043	-0.029
	(0.048)	(0.049)	(0.065)
Constant	-0.117*	-0.245***	-0.602**
	(0.063)	(0.059)	(0.296)
UC FEs	Yes	Yes	Yes
R-Squared	0.294	0.260	0.142
# Observations	2435	2435	2384
	Panel B: Men's Interest		
T1only	-0.018	0.075	-0.027
	(0.067)	(0.072)	(0.070)
T2only	-0.085	0.073	-0.020
	(0.067)	(0.069)	(0.076)
T1+T2	0.049	0.182**	0.085
	(0.068)	(0.075)	(0.076)
Within T Control	0.014	0.069	0.029
	(0.061)	(0.066)	(0.067)
Constant	0.350*	0.836***	0.215
	(0.185)	(0.106)	(0.173)
UC FEs	Yes	Yes	Yes
R-Squared	0.253	0.194	0.227
# Observations	2434	2434	2413

Notes: All specifications show results using OLS estimation and employ strata (Union Council) fixed effects. Standard errors in parentheses are clustered at the ward level. All three outcomes are standardized. All three outcomes used in this table are combined into a standardized index to form the outcome variable for Column (2) of Table C.1. Column (1) uses responses to the question "How interested are you in political TV shows?" as outcome. Column (2) uses responses to the question "How interested are you in political issues / topics or problems?". Column (3) uses responses to the question "How interested would you say you were in the 2018 Election?". All questions are asked on a Likert scale. * p<0.10, ** p<0.05, *** p<0.01

Table C.6: Treatment Effects on Self-Efficacy

	Panel A: Women's Responses'		
	(1) **Qualified to Participate**	(2) **Informed about Voting**	(3) **Politics too Complicated**
T1only	-0.002	-0.025	-0.012
	(0.075)	(0.078)	(0.071)
T2only	-0.037	0.011	0.019
	(0.080)	(0.086)	(0.076)
T1+T2	-0.020	-0.015	0.012
	(0.070)	(0.081)	(0.073)
Within T Control	0.041	0.124	-0.076
	(0.066)	(0.076)	(0.067)
Constant	-0.788***	-0.309	0.377
	(0.271)	(0.332)	(0.240)
UC FEs	Yes	Yes	Yes
R-Squared	0.227	0.087	0.133
# Observations	2411	2363	2410
	Panel B: Men's Responses		
T1only	-0.030	0.075	-0.014
	(0.074)	(0.065)	(0.077)
T2only	-0.064	0.077	-0.063
	(0.080)	(0.073)	(0.078)
T1+T2	0.061	0.051	-0.012
	(0.080)	(0.073)	(0.083)
Within T Control	-0.115*	0.039	-0.041
	(0.069)	(0.064)	(0.071)
Constant	-0.393	-0.361	0.381
	(0.318)	(0.416)	(0.348)
UC FEs	Yes	Yes	Yes
R-Squared	0.258	0.157	0.147
# Observations	2428	2415	2427

Notes: All specifications show results using OLS estimation and employ strata (Union Council) fixed effects. Standard errors in parentheses are clustered at the ward level. All three outcomes are standardized. All three outcomes used in this table are combined into a standardized index to form the outcome variable for Column (3) of Table C.1. The questions used as outcomes are agreement on a Likert scale with the following statements: Column (1) I consider myself well-qualified to participate in politics as a citizen, Column (2) I think that I am well-informed about the process of how to cast my vote in the next election and (disagreement with) Column (3) Sometimes politics and government seem so complicated that a person like me can't really understand what's going on. * p<0.10, ** p<0.05, *** p<0.01

Table C.7: Treatment Effects on Views on Men's Restrictions

	Panel A: Women's Views on Restrictions if:			
	Vote Differently	**Long Lines**	**Threat of Violence**	**Interferes w/ HH Duties**
T1only	0.105	0.010	0.044	-0.016
	(0.066)	(0.065)	(0.074)	(0.074)
T2only	0.079	-0.029	-0.039	-0.035
	(0.070)	(0.070)	(0.080)	(0.078)
T1+T2	0.087	0.044	0.083	0.046
	(0.068)	(0.067)	(0.074)	(0.074)
Within T Control	0.059	0.061	-0.018	-0.002
	(0.059)	(0.058)	(0.069)	(0.068)
Constant	0.726***	0.382*	0.096	0.018
	(0.128)	(0.219)	(0.365)	(0.366)
UC FEs	Yes	Yes	Yes	Yes
R-Squared	0.357	0.302	0.235	0.224
# Observations	2426	2428	2409	2427
	Panel B: Men's Views on Restrictions if:			
T1only	0.129	0.016	-0.026	0.042
	(0.081)	(0.078)	(0.072)	(0.072)
T2only	0.162*	0.056	-0.104	0.085
	(0.084)	(0.082)	(0.079)	(0.078)
T1+T2	-0.041	-0.040	-0.127*	0.003
	(0.079)	(0.082)	(0.072)	(0.073)
Within T Control	-0.017	0.014	-0.094	0.003
	(0.071)	(0.068)	(0.065)	(0.068)
Constant	0.096	-0.396	0.358	0.015
	(0.345)	(0.445)	(0.255)	(0.302)
UC FEs	Yes	Yes	Yes	Yes
R-Squared	0.249	0.210	0.231	0.211
# Observations	2433	2433	2432	2432

Notes: All specifications show results using OLS estimation and employ strata (Union Council) fixed effects. Standard errors in parentheses are clustered at the ward level. All three outcomes used in this table are indicator variables combined into a standardized index to form the outcome variable for Column (4) of Table C.1. Each column uses answers to questions about conditions under which the respondent thinks it is appropriate for men to stop women from voting. In Column (1), the condition is "They (men) think women will vote for a different candidate/party than the one they support". In Column (2), the condition is "The lines are expected to be very long and women might have to stand outside the polling station while waiting". In Column (3), the condition is "There is a chance of fights breaking out at the polling station". In Column (4), the condition is "They (men) think it will interfere with women's household duties". * p<0.10, ** p<0.05, *** p<0.01.

Table C.8: Treatment Effects on Men's Election Day Support

	Panel A: Women's Responses		
	(1) **Organizing Transport**	(2) **Sharing HH Chores**	(3) **Waiting at PS**
T1only	-0.086	0.044	-0.123
	(0.070)	(0.068)	(0.076)
T2only	-0.023	-0.025	-0.023
	(0.070)	(0.077)	(0.082)
T1+T2	0.132*	0.116	0.113
	(0.075)	(0.072)	(0.081)
Within T Control	0.057	0.058	0.091
	(0.062)	(0.063)	(0.070)
Constant	0.183	-0.242	0.527***
	(0.125)	(0.235)	(0.117)
UC FEs	Yes	Yes	Yes
R-Squared	0.224	0.240	0.161
# Observations	2374	2377	2372
	Panel B: Men's Responses		
T1only	0.051	-0.036	0.107
	(0.070)	(0.072)	(0.076)
T2only	0.040	-0.143**	0.116
	(0.079)	(0.072)	(0.079)
T1+T2	0.152**	0.089	0.173**
	(0.076)	(0.077)	(0.079)
Within T Control	0.068	0.006	0.108
	(0.068)	(0.067)	(0.070)
Constant	0.668***	0.592**	0.491**
	(0.137)	(0.264)	(0.191)
UC FEs	Yes	Yes	Yes
R-Squared	0.168	0.186	0.137
# Observations	2418	2394	2359

Notes: All specifications show results using OLS estimation and employ strata (Union Council) fixed effects. Standard errors in parentheses are clustered at the ward level. All three outcomes used in this table are indicator variables combined into a standardized index to form the outcome variable for Column (5) of Table C.1. For women (Panel A), the questions used as outcomes are responses to the question "How willing were the men in your household to help with the following things before the election/on election day?". For men (Panel B), the questions used as outcome are yes or no responses to the question "Did you do any of the following before the election/on election day?". The relevant actions for each column respectively are (1) Organizing transport/taking women to the polling station on election day, (2) Sharing household duties so that women had time to vote and (3) Waiting for women at the polling station. * p<0.10, ** p<0.05, *** p<0.01.

D Robustness Checks

D.1 Women's Turnout with Controls

Table D.1: Results: Women's Turnout (ITT)

	Women's Turnout – Household Level Controls			
	(1)	(2)	(3)	(4)
	HH Proportion	**HH Proportion**	**HH Proportion**	**HH Proportion**
T1 Only:Women Canvassed	0.013			
	(0.029)			
T2 Only:Men Canvassed	0.055*			
	(0.031)			
T1+T2: Women and Men Both	0.079**			
	(0.032)			
T1: Women Canvassed		0.018		0.013
		(0.020)		(0.029)
T2: Men Canvassed			0.060**	0.055*
			(0.024)	(0.031)
T1*T2				0.011
				(0.040)
Within T Control	0.024	0.010	0.019	0.024
	(0.028)	(0.026)	(0.026)	(0.028)
Constant	0.538***	0.550***	0.542***	0.538***
	(0.029)	(0.026)	(0.026)	(0.029)
R-Squared	0.156	0.153	0.155	0.156
# Observations	2146	2146	2146	2146
P-Value: T1only=T2only	0.185			
P-Value: T1only=T1+T2	0.038			
P-Value: T2only=T1+T2	0.374			

Notes: All specifications show results using OLS estimation, include block (Union Council) fixed effects and control for individual level randomizations. Standard errors in parentheses are clustered at the ward level. The outcome variable is women's turnout at the household level calculated as the number of women who voted (as verified by thumb ink marks) as a proportion of women who have an identity card and are therefore eligible to vote. All models include controls at the household level for the total number of adult men and women in the household, whether the household has a joint (vs. nuclear) family, the presence of young children and elderly members who require care, and a standardized index of assets. *p<0.10, ** p<0.05, *** p<0.01

D.2 Men's Turnout

In Table D.2 we report results on male turnout using the same specifications as in Table 5 of the main paper. We do not find any evidence of effects of any of the treatment conditions on male turnout across specifications.

The outcome measure for this analysis is the number of men in a household for whom turnout could be visually verified (measured during the turnout verification exercise) divided by the number of male household members eligible to vote (measured at baseline). However, due to the time constraint during the turnout verification exercise (see paper section Outcome Data: Turnout), we could reach far fewer men per household than women to verify turnout. This is reflected in the low mean in the control group (0.34, sd 0.36) which is far lower than official male turnout numbers in these constituencies which ranges from 57% to 63%, and is indicative of overall measurement error in this outcome.

Table D.2: Results: Men's Turnout (ITT)

	Men's Turnout – Unadjusted			
	(1)	(2)	(3)	(4)
	HH Proportion	**HH Proportion**	**HH Proportion**	**HH Proportion**
T1 Only:Women Canvassed	0.013			
	(0.026)			
T2 Only:Men Canvassed	0.005			
	(0.027)			
T1+T2: Women and Men Both	0.035			
	(0.028)			
T1: Women Canvassed		0.022		0.013
		(0.018)		(0.026)
T2: Men Canvassed			0.013	0.005
			(0.021)	(0.027)
T1*T2				0.017
				(0.036)
Within T Control	0.024	0.024	0.019	0.024
	(0.025)	(0.023)	(0.022)	(0.025)
Constant	0.341***	0.341***	0.346***	0.341***
	(0.016)	(0.013)	(0.013)	(0.016)
R-Squared	0.200	0.199	0.199	0.200
# Observations	2190	2190	2190	2190
P-Value: T1only=T2only	0.758			
P-Value: T1only=T1+T2	0.442			
P-Value: T2only=T1+T2	0.218			

Notes: All specifications show results using OLS estimation, include block (Union Council) fixed effects and control for individual level randomizations. Standard errors in parentheses are clustered at the ward level. The outcome variable is men's turnout at the household level calculated as the number of men who voted (as verified by thumb ink marks) as a proportion of men who have an identity card and are therefore eligible to vote. This table shows unadjusted results, *p<0.10, ** p<0.05, *** p<0.01

D.3 Spillovers to Untreated Households

To be able to detect spillovers within clusters, we implemented a partial population design (Baird et al., 2016), whereby in the second stage of randomization, 4 of 5 study households in each treatment cluster were randomly assigned to receive the treatment, and 1 remaining household designated as a "Within-Treatment Control". In all main analyses reported in the paper we report the coefficient on the pooled "Within Treatment Controls" across treatment conditions. Table D.3 examines the spillover effect on Within treatment control households for each treatment condition separately. We do not find any significant differences, however we would caveat that we are not particularly well powered for these comparisons.

Table D.3: Effects on Untreated Households in Treatment Clusters

	(1)
	HH Proportion
T1 Only:Women Canvassed	0.011
	(0.028)
T2 Only:Men Canvassed	0.054*
	(0.031)
T1+T2: Women and Men Both	0.081**
	(0.032)
Within T1 Control	-0.032
	(0.041)
Within T2 Control	0.050
	(0.041)
Within T1+T2 Control	0.046
	(0.043)
Constant	0.562***
	(0.017)
R-Squared	0.154
# Observations	2149
P-Value: T1-Control=T2-Control	0.117
P-Value: T1-Control=T1+T2-Control	0.148
P-Value: T2-Control=T1+T2-Control	0.951

Notes: All specifications show results using OLS estimation, include block (Union Council) fixed effects and control for individual level randomizations. Standard errors in parentheses are clustered at the ward level. The outcome variable is women's turnout at the household level calculated as the number of women who voted (as verified by thumb ink marks) as a proportion of women who have an identity card and are therefore eligible to vote. *p<0.10, ** p<0.05, *** p<0.01

D.4 Compliance

The paper section titled "Recall and Compliance" describes our approach towards measuring compliance. Table D.4 shows results from a regression of our compliance measure on treatment status. Compliance is not significantly different between T1 and T2, however it is slightly lower in T1+T2 due to the higher bar for compliance (completion of 2 successful visits targeted to women and men respectively).

Table D.4: Proportion of Compliers, by Treatment Status

	(1) **Compliance**
T1 Only:Women Canvassed	0.968***
	(0.009)
T2 Only:Men Canvassed	0.965***
	(0.011)
T1+T2: Women and Men Both	0.945***
	(0.012)
Within T Control	0.002
	(0.005)
Constant	0.001
	(0.004)
R-Squared	0.908
# Observations	2390
P-Value: T1only=T2only	0.787
P-Value: T1only=T1+T2	0.094
P-Value: T2only=T1+T2	0.194

Notes: The regression uses OLS estimation and employs block (Union Council) fixed effects. The outcome is a binary indicator for whether a canvasser could successfully deliver the intervention within 3 attempts at contact * p<0.10, ** p<0.05, *** p<0.01

We estimate the complier average causal effect (CACE) for our preferred specification (Table 5, column 1). To do this we measure compliance as described above, and employ an instrumental variable approach where random assignment to treatment is used as an instrument for treatment status in a model following our preferred specification. Table D.5 reports results. This analysis relies on the following assumptions:

1. The treatment is randomly assigned
2. There is a positive share of compilers (met since the compliance rate is 96%)
3. Monotonicity, i.e. assignment to treatment does not make one less likely to be treated (met since the coefficients on a model regressing compliance on treatment status are all positive and significant (Table D.4)
4. Exclusion restriction: individuals respond to the treatment, not treatment assignment (plausibly met because individuals are not aware of treatment status)

The CACE estimates are very similar to the ITT estimates in Table 5 (Column 1).

Table D.5: Complier Average Causal Effect

	(1) **Women's Turnout – CACE**
T1 Only:Women Canvassed	0.012
	(0.028)
T2 Only:Men Canvassed	0.056*
	(0.032)
T1+T2: Women and Men Both	0.085***
	(0.033)
Within T Control	0.022
	(0.027)
Constant	0.725***
	(0.091)
R-Squared	0.154
# Observations	2149
P-Value: T1only=T2only	0.163
P-Value: T1only=T1+T2	0.022
P-Value: T2only=T1+T2	0.293

Notes: The regression uses OLS estimation and employs block (Union Council) fixed effects. Random assignment to each of 3 experimental treatment conditions is used as an instrument for compliance with the treatment conditions * p<0.10, ** p<0.05, *** p<0.01

D.5 Attrition

We analyze attrition in our sample between baseline and turnout verification to assess whether it is correlated with treatment status. First, we define a dummy variable that indicates whether a household was surveyed at baseline but could not be reached for turnout verification in the days following the election. In Table D.6 we report results from a regression of this attrition dummy on indicators of treatment status and do not find evidence that assignment to treatment is significantly correlated with the probability of attrition; there is also no evidence of significant differences in the probability of attrition between different treatment groups.

Table D.6: Probability of Attrition, by Treatment Status

	(1)
	Attrition
T1 Only:Women Canvassed	-0.015
	(0.021)
T2 Only:Men Canvassed	-0.002
	(0.021)
T1+T2: Women and Men Both	-0.008
	(0.021)
Within T Control	0.000
	(0.022)
Constant	0.146***
	(0.014)
R-Squared	0.078
# Observations	2500
P-Value: T1only=T2only	0.562
P-Value: T1only=T1+T2	0.755
P-Value: T2only=T1+T2	0.788

Notes: The regression uses OLS estimation and employs block (Union Council) fixed effects. The outcome is a binary indicator for whether a household was observed at baseline but not at the time of turnout verification * p<0.10, ** p<0.05, *** p<0.01

Nevertheless, to account for the possibility nonrandom attrition in the sample, we follow (Lee, 2009) and report the lower and upper bounds on our main treatment effects using trimming bounds. We report these results in Table D.7, Columns 2-3 and find only minimal differences from the original estimation.

Table D.7: Lee Trimming Bounds for Treatment Effect on Turnout

	(1) **No Correction**	(2) **Lower Bound**	(3) **Upper Bound**
T1 Only:Women Canvassed	0.012	0.001	0.025
	(0.028)	(0.029)	(0.028)
T2 Only:Men Canvassed	0.054*	0.053*	0.058*
	(0.031)	(0.031)	(0.031)
T1+T2: Women and Men Both	0.080**	0.074**	0.090***
	(0.032)	(0.032)	(0.032)
Within T Control	0.022	0.017	0.026
	(0.028)	(0.028)	(0.028)
Constant	0.562***	0.564***	0.561***
	(0.017)	(0.018)	(0.018)
R-Squared	0.153	0.157	0.158
# Observations	2149	2125	2125

Notes: The regression uses OLS estimation and employs block (Union Council) fixed effects and control for individual level randomizations. Standard errors in parentheses are clustered at the ward level. The outcome variable is women's turnout at the household level calculated as the number of women who voted (as verified by thumb ink marks) as a proportion of women who have an identity card and are therefore eligible to vote. * p<0.10, ** p<0.05, *** p<0.01

D.6 Baseline Survey With Political Content

Table D.8 shows our main results on turnout, controlling for an indicator for whether the baseline respondent from a household was randomized into receiving a survey that included questions about political participation and the upcoming elections. We do not find that the baseline survey affected participation, and there is no change in the estimated treatment effects when we control for this.

Table D.8: Results: Women's Turnout (ITT) Controlling for Baseline Political Survey

	(1)
	HH Proportion
T1 Only:Women Canvassed	0.012
	(0.028)
T2 Only:Men Canvassed	0.054*
	(0.031)
T1+T2: Women and Men Both	0.080**
	(0.032)
Within T Control	0.022
	(0.028)
Political Survey	-0.005
	(0.021)
Constant	0.571***
	(0.041)
R-Squared	0.153
# Observations	2149

Notes: The regression uses OLS estimation and employs block (Union Council) fixed effects and control for individual level randomizations. Standard errors in parentheses are clustered at the ward level. The outcome variable is women's turnout at the household level calculated as the number of women who voted (as verified by thumb ink marks) as a proportion of women who have an identity card and are therefore eligible to vote. * p<0.10, ** p<0.05, *** p<0.01

D.7 Heterogeneous Effects by Number of Individuals Canvassed

To examine whether we observe higher treatment effects in the T1+T2 condition because we treated a higher number of individuals in that condition, we analyze treatment effects by the number of individuals treated within each treatment arm. In particular, we test whether treating more than the minimum prescribed number of individuals is associated with a higher treatment effect. For T1 and T2, the minimum prescribed number is 1, while in T1+T2, the minimum prescribed number is 2. The mobilizers aimed to treat every household member of the relevant gender who was present and available, which meant that in many cases we treated more than the minimum prescribed number. We find that treatment effects are not significantly higher when we treat more individuals. On the contrary, the treatment effect co-efficient for T2 is lower when more than 1 men is treated, although this difference is not statistically significant. This analysis suggests that the higher treatment effects in the T1+T2 condition are not driven by a difference in the number of treated individuals.

Table D.9: Results: Treatment Effects by Number of Individuals Treated

	(1)
	HH Proportion
T1only: 1 Woman Treated	0.015
	(0.031)
T1only: 2+ Women Treated	0.006
	(0.039)
T2only: 1 Man Treated	0.065**
	(0.033)
T2only: 2+Men Treated	0.010
	(0.049)
T1+T2: 2 People Treated	0.071*
	(0.036)
T1+T2: 3+ People Treated	0.100***
	(0.036)
Within T Control	0.023
	(0.028)
Constant	0.561***
	(0.017)
R-Squared	0.154
# Observations	2149.000
P-Value: T1: 1 vs. 2+ Treated	0.820
P-Value: T2: 1 vs. 2+ Treated	0.270
P-Value: T3: 2 vs. 3+ Treated	0.434

Notes: The regression uses OLS estimation and employ block (Union Council) fixed effects and control for individual level randomizations. Standard errors in parentheses are clustered at the ward level. The outcome is the proportion of women in the HH who voted. For each of T1, T2 and T1+T2, we replace the treatment indicator with two indicators each: one for the minimum prescribed number of individuals to be treated (1 for T1 and T2; 2 for T1+T2) and another for any value higher than the minimum prescribed.* p<0.10, ** p<0.05, *** p<0.01

D.8 Men's Supportive Behavior (Alternate Specification)

In the paper, we report results on the effect of treatment on men's supportive behavior using a set of difference-in-difference estimates which allow for clearer presentation. Here, we estimate the effect of the treatments using a specification analogous to the one used for our main turnout results (Table 5 Column 3 in the main paper), and show that the results are similar both in terms of substance and statistical significance. We use the following specification:

$$Y_i = \beta_1 WS_i + \beta_{2i} Treatment_i + \beta_{3i} WS * Treatment_i + \delta_i + \gamma_s \quad (1)$$

where WS_i is an indicator for whether the sticker offered to the male respondent in study household i was one with a message supporting women's role in democracy, $Treatment_i$ denotes separate indicators for T1 only, T2 only, T1+T2, and within treatment control. γ_s are union council fixed effects. Y_i is an indicator for whether the male respondent agreed to having the sticker placed on the entry-way to his residence. Standard errors are clustered at the ward level, which is the level of randomization. The co-efficient of interest is β_3, which estimates whether treatment affects men's relative propensity to accept the sticker with a supportive message about women's participation versus the generic support of democracy sticker.

The results are shown in Table D.10. They key result is the coefficient on the $(T1+T2)*WS$ term, which shows that the relative take-up of the women's support sticker was 7.6 percentage points higher among men in households that received visits targeted to both women and men (T1+T2), compared to those in the control group.

Table D.10: Results: Men's Support for Women's Role in Democracy (ITT; Behavioral Measure)

	Men's Support for Women's Role in Democracy
	(1)
	Binary Takeup
Support for Women Message (WS)	-0.047**
	(0.022)
T1 Only:Women Canvassed	0.006
	(0.021)
T2 Only:Men Canvassed	-0.017
	(0.023)
T1+T2: Women and Men Both	-0.031
	(0.022)
Within T Control	-0.021
	(0.023)
T1only*WS	-0.013
	(0.031)
T2only*WS	0.027
	(0.031)
(T1+T2)*WS	0.075***
	(0.028)
Within T Control*WS	0.015
	(0.036)
Constant	0.949***
	(0.014)
# Observations	2434
P-Value: T1only*WS=T2only*WS	0.206
P-Value: T1only*WS=(T1+T2)*WS	0.002
P-Value: T2only*WS=(T1+T2)*WS	0.100

Notes: All specifications show results using OLS estimation, include block (Union Council) fixed effects and control for individual level randomizations. Standard errors in parentheses are clustered at the ward level. The outcome is an indicator for whether the male respondent agreed to post the offered sticker on the entry-way to their residence. WS indicates whether the sticker offered to them had a message indicating support for women's role in democracy. * p<0.10, ** p<0.05, *** p<0.01

References

Anderson, Michael L. 2008. "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American statistical Association* 103(484):1481–1495.

Baird, Sarah, J Aislinn Bohren, Craig McIntosh and Berk Özler. 2016. "Optimal design of experiments in the presence of interference." *Review of Economics and Statistics* (0).

Giné, Xavier and Ghazala Mansuri. 2018. "Together we will: Experimental evidence on female voting behavior in Pakistan." *American Economic Journal: Applied Economics* 10(1):207–35.

Lee, David S. 2009. "Training, wages, and sample selection: Estimating sharp bounds on treatment effects." *The Review of Economic Studies* 76(3):1071–1102.

Slough, Tara. 2020. "The Ethics of Electoral Experimentation: Design-Based Recommendations.".