Codebook for

Can Party Elites Shape the Rank-and-File? Evidence from a Recruitment Campaign in India

By Saad Gulzar, Durgesh Pathak, Sarah Thompson, Aliz Tóth

General Notes:

There is one main dataset included in this replication package:

1. aap_pamphlets_replication.csv is the main dataset. The unit of observation is the pamphlet. It includes information on the treatment status of the pamphlet and whether the pamphlet has received a call back from potential volunteers as well as the characteristics of potential volunteers. Note that all main outcomes have been multiplied by a 1,000 to make the results more easily interpretable. This dataset is created from “compliance.csv”, “onboarding_survey.csv” and “recall_survey.csv”. For more details, see below.

In addition, we also include six auxiliary datasets for additional analysis:

2. compliance.csv is a dataset that includes details about mid-level members' (vice presidents) treatment assignment and which kind of pamphlets they have received.

3. onboarding_survey.csv is a dataset of survey respondents that AAP was able to contact after the recruitment campaign.

4. recall_survey.csv is a dataset of respondents who were recontacted by AAP in the long-term retention survey.

5. compliance _timeline.csv is a dataset that includes details about the timeline of pamphlet distribution for each mid-level party member.

6. treatment_assignment.csv is a dataset of treatment assignment for each assembly constituency. Column names are assembly constituency names. Each row represents the treatment assignment for a mid-level party member. The numeric IDs of treatments in each row were used on sign-up sheets for mid-level party members to assign them to different treatment pamphlets.

7. vpsurvey.csv includes information from a survey with mid-level party members.

8. census_2011.csv is a dataset of all villages in Jharkhand with demographic information drawn from the Census of India 2011

9. baseline_observation.csv is a dataset of responses from the ''missed call survey'' with potential volunteers. The survey only includes respondents who called a number associated with the baseline pamphlet.

Main dataset (aap_ pamphlets_ replication.csv)

Variable name	Explanation
id_vp2	numeric ID for the mid-level member distributing pamphlets
treatment_assigned	the type of treatment the pamphlet was assigned to: 1 - candidacy (M), 2 - career (M), 3 - ideology (M), 4 - policy (M), 5 - baseline (M), 6 - candidacy (F), 7 - career (F), 8 - ideology (F), 9 - policy (F), 10 - baseline (F)
fem_treatment_assigned	whether female treatment was assigned to the pamphlet (1 - yes, 0 - no)
t_base	whether baseline treatment was assigned to the pamphlet (1 - yes, 0 - no)
assembly_id	numeric ID for the assembly constituency where the pamphlet was distributed
missedcall	whether the pamphlet received a call back from a potential volunteer (1000 - yes, 0 - no)
consented_21	whether the pamphlet received a call back from a volunteer who agreed to be interviewed again in 2021 by the party (1000 - yes, 0 - no)
excluded_group	whether the pamphlet received a call back from a potential volunteer who is a member of an excluded group: excluded caste/tribe, religion, or gender (1000 - yes, 0 - no)
t_fem	pamphlet assigned to female treatment (1- yes, 0 - no)
t_cand	pamphlet assigned to candidacy treatment (1 - yes, 0 - no)
t_poli	pamphlet assigned to policy treatment (1 - yes, 0 - no)
t_ideo	pamphlet assigned to ideology treatment (1 - yes, 0 - no)
t_care	pamphlet assigned to career treatment (1 - yes, 0 - no)
t_fem_takeup	pamphlet received female treatment (1 - yes, 0 - no)
t_base_male	pamphlet assigned to male baseline treatment (1 - yes, 0 - no)
t_base_female	pamphlet assigned to female baseline treatment (1 - yes, 0 - no)
t_base_male_takeup	pamphlet received male baseline treatment (1 - yes, 0 - no)
t_base_takeup	pamphlet received baseline treatment (1 - yes, 0 - no)
t_base_female_takeup	pamphlet received female baseline treatment (1 - yes, 0 - no)
t_cand_male	pamphlet assigned to male candidacy treatment (1 - yes, 0 - no)
t_cand_female	pamphlet assigned to female candidacy treatment (1 - yes, 0 - no)
t_cand_male_takeup	pamphlet received male candidacy treatment (1 - yes, 0 - no)

t_cand_takeup	pamphlet received candidacy treatment (1 - yes, 0 - no)
t_cand_female_takeup	pamphlet received female candidacy treatment (1 - yes, 0 - no)
t_care_male	pamphlet assigned male career treatment (1 - yes, 0 - no)
t_care_female	pamphlet assigned female career treatment (1 - yes, 0 - no)
t_care_male_takeup	pamphlet received male career treatment (1 - yes, 0 - no)
t_care_takeup	pamphlet received career treatment (1 - yes, 0 - no)
t_care_female_takeup	pamphlet received female career treatment (1 - yes, 0 - no)
t_ideo_male	pamphlet assigned male ideology treatment (1 - yes, 0 - no)
t_ideo_female	pamphlet assigned female ideology treatment (1 - yes, 0 - no)
t_ideo_male_takeup	pamphlet received male ideology treatment (1 - yes, 0 - no)
t_ideo_takeup	pamphlet received ideology treatment (1 - yes, 0 - no)
t_ideo_female_takeup	pamphlet received female ideology treatment (1 - yes, 0 - no)
t_poli_male	pamphlet assigned male policy treatment (1 - yes, 0 - no)
t_poli_female	pamphlet assigned female policy treatment (1 - yes, 0 - no)
t_poli_male_takeup	pamphlet received male policy treatment (1 - yes, 0 - no)
t_poli_takeup	pamphlet received policy treatment (1 - yes, 0 - no)
t_poli_female_takeup	pamphlet received female policy treatment (1 - yes, 0 - no)
female	pamphlet received callback from a potential volunteer who is a woman (1000 - yes, 0 - no)
excluded_caste	pamphlet received callback from a potential volunteer who is a member of an excluded caste or tribe (Scheduled Caste / Scheduled Tribe / Other Backward Classes) (1000 - yes, 0 - no)
sc_caste	pamphlet received callback from a potential volunteer who is a member of Scheduled Castes (1000 - yes, 0 - no)
st_group	pamphlet received callback from a potential volunteer who is a member of Scheduled Tribes (1000 - yes, 0 - no)
obc_caste	pamphlet received a callback from a potential volunteer who is a member of Other Backward Classes (1000 - yes, 0 - no)
excluded_religion	pamphlet received a callback from a potential volunteer who is a member of an excluded religion (non- Hindu) (1000 - yes, 0 - no)
muslim	pamphlet received a callback from a potential volunteer who is a Muslim (1000 - yes, 0 - no)

christian	pamphlet received a callback from a potential volunteer who is a Christian (1000 - yes, 0 - no)
other_relig	pamphlet received a callback from a potential volunteer who is of any other religion (1000 - yes, 0 - no)
hindu	pamphlet received a callback from a potential volunteer who is a Hindu (1000 - yes, 0 - no)
skill_index	pamphlet received a callback from a potential volunteer who is skilled: either employed, has at least high school education, voted in prior elections, or has prior campaign volunteering experience (1000 - yes, 0 - no)
any_employ	pamphlet received a callback from a potential volunteer who is employed (1000 - yes, 0 - no)
educ_hi	pamphlet received a callback from a potential volunteer who has at least high school education (1000 - yes, 0 - no)
prior_vote	pamphlet received a callback from a potential volunteer who has voted in prior elections (1000 - yes, 0 - no)
prior_volunteer	pamphlet received a callback from a potential volunteer who has prior volunteering experience (1000 - yes, 0 - no)
excluded_group_21	pamphlet received a callback from a potential volunteer who is a member of an excluded group and was re-contacted by AAP in 2021 (1000 - yes, 0 - no)
skill_index_21	pamphlet received a callback from a potential volunteer who is skilled and was re-contacted by AAP in 2021 (1000 - yes, 0 - no)
newmemb	pamphlet received a callback from a potential volunteer who indicated their geographic location (1000 - yes, 0 - no)
motivmemb	pamphlet received a callback from a potential volunteer who agreed to volunteer for AAP (1000 - yes, 0 - no)
included_group	pamphlet received a callback from a potential volunteer who is a member of an included group (not excluded group) (1000 - yes, 0 - no)
unskilled	pamphlet received a callback from a potential volunteer who is low skilled: no higher education, employment, prior voting or volunteering experience (1000 - yes, 0 - no)
female_21	pamphlet received a callback from a potential volunteer who is a woman and was re-contacted by AAP in 2021 (1000 - yes, 0 - no)
excluded_caste_21	pamphlet received a callback from a potential volunteer who is a member of an excluded caste or tribe and was re-contacted by AAP in 2021 (1000 - yes, 0 - no)
excluded_religion_21	pamphlet received a callback from a potential volunteer who is a member of an excluded religion and was re-contacted by AAP in 2021 (1000 - yes, 0 - no)

any_employ_21	pamphlet received a callback from a potential volunteer who is employed and was re-contacted by AAP in 2021 (1000 - yes, 0 - no)
educ_hi_21	pamphlet received a callback from a potential volunteer who has at least high school education and was re-contacted by AAP in 2021 (1000 - yes, 0 - no)
prior_vote_21	pamphlet received a callback from a potential volunteer who has voted in prior elections and was re-contacted by AAP in 2021 (1000 - yes, 0 - no)
prior_volunteer_21	pamphlet received a callback from a potential volunteer who has prior volunteering experience and was re-contacted by AAP in 2021 (1000 - yes, 0 - no)

Compliance Dataset 1 (compliance.csv)

Variable name	Explanation
Assembly_name	name of the constituency assembly where a mid-level member (vice president) was assigned to work
Assembly_ID	numerical ID for each constituency
ID	numerical ID of the mid-level member (vice president)
name_order	numerical ID for sorting observations in the alphabetical order of mid-level member names
Treatment_assigned	the type of pamphlet assigned to the mid-level member (vice president): 1 - candidacy (M), 2 - career (M), 3 - ideology (M), 4 - policy (M), 5 - baseline (M), 6 - candidacy (F), 7 - career (F), 8 - ideology (F), 9 - policy (F), 10 - baseline (F).
Pamphlet_ID_final	the type of pamphlet received by the mid-level member (vice president): 1 - candidacy (M), 2 - career (M), 3 - ideology (M), 4 - policy (M), 5 - baseline (M), 6 - candidacy (F), 7 - career (F), 8 - ideology (F), 9 - policy (F), 10 - baseline (F).
GP_Ward_final	the final area assigned to the mid-level member to disribute pamphlets in
geo_unit	whether the area assigned to the mid-level member is a rural gram panchayat (village cluster) - GP, an urban ward, sector, or residential colony, or a polling station booth
Pamphlet_received	number of pamphlets received by the mid-level party member to distribute

Compliance Dataset 2 (compliance_timeline.csv)

Variable name	Explanation
Assembly_name	name of the constituency assembly where a mid-level member (vice president) was assigned to work
ID	numerical ID of the mid-level member (vice president)
Treatment_assigned	the type of pamphlet assigned to the mid-level member (vice president): 1 - candidacy (M), 2 - career (M), 3 - ideology (M), 4 - policy (M), 5 - baseline (M), 6 - candidacy (F), 7 - career (F), 8 - ideology (F), 9 - policy (F), 10 - baseline (F).
Pamphlet_ID_final	the type of pamphlet received by the mid-level member (vice president): 1 - candidacy (M), 2 - career (M), 3 - ideology (M), 4 - policy (M), 5 - baseline (M), 6 - candidacy (F), 7 - career (F), 8 - ideology (F), 9 - policy (F), 10 - baseline (F).
Pamphlet_received	total number of pamphlets received by the mid-level party member
date1	first date when the party checked in with mid-level party members about the number of pamphlets distributed
date2	second date when the party checked in with mid-level party members about the number of pamphlets distributed
date3	third date when the party checked in with mid-level party members about the number of pamphlets distributed
date4	fourth date when the party checked in with mid-level party members about the number of pamphlets distributed
date5	fifth date when the party checked in with mid-level party members about the number of pamphlets distributed
date6	sixth date when the party checked in with mid-level party members about the number of pamphlets distributed
distributed1	number of pamphlets distributed by the first check-in date
distributed2	number of pamphlets distributed by the second check-in date
distributed3	number of pamphlets distributed by the third check-in date
distributed4	number of pamphlets distributed by the fourth check-in date
distributed5	number of pamphlets distributed by the fifth check-in date
distributed6	number of pamphlets distributed by the sixth check-in date

VP Survey (vpsurvey.csv)

Variable name	Explanation
vp_gender	mid-level party member's gender, Male=1 ; Female= 0; Others= 9
vp_caste	mid-level party member's caste (General - general caste; SC - Scheduled Caste, ST - Scheduled Tribe, OBC - Other Backward Classes, "I do not associate with any caste", "I would prefer not to answer")
vp_religion	mid-level party member's religion
vp_education	mid-level party member's education level

Missed Call Survey (onboarding_survey.csv)

Variable name	Explanation
uniq_id	unique, numerical ID for each respondent
unique_id_recall	unique ID linking onboarding survey respondent to responses in the long-term retention survey
pamph_ID_treat	ID of the pamphlet that the call has been matched to
baseline	whether or not this call was recruited using a status quo pamphlet (0 = no, 1 = yes)
gender	gender of caller (Male = 1; Female = 0; Refused to answer=99)
caste	caste of caller
caste_others	other caste affiliation
religion	religion of caller
religion_others	other religious affiliation
education	education of caller
education_others	other educational qualification of the caller
literate	if the caller did not attend formal schooling, can the caller read or write (1 - yes, 0 - no)
district_mcs	administrative district that the respondent reported calling from
assembly_mcs	constituency from which the respondent reported calling
gpname	gram panchayat from which the respondent reported calling
employed	whether respondent was employed or not (1 – yes, 0 – no)
vote_2014	whether or not the respondent voted in the 2014 elections (1 – yes, 0 – no)
volunteer_party	whether or not the respondent has ever volunteered for a political party (1 – yes, 0 – no)
time_aap	whether respondent wanted to volunteer for AAP (full time=1; part time= 2; no time= 0; refused to answer= 99; others= 00)

Missed Call Survey Baseline Observations (baseline_observations.csv)

Variable name	Explanation
gender	gender of caller (Male = 1; Female = 0; Refused to answer=99)
caste	caste of caller
religion	religion of caller
education	education of caller
education_others	other educational qualification of the caller
literate	if the caller did not attend formal schooling, can the caller read or write (1 - yes, 0 - no)

Long-term Retention Survey (recall_survey.csv)

Variable name	Explanation
unique_id_recall	unique ID for linking respondents to onboarding survey respondents
contacted_21	whether AAP succeeded in contacting respondent in 2021
consented_21	whether AAP consented to taking the retention survey in 2021

Jharkhand Census (census_2011.csv)

Variable names	Explanation
TOT_POP	total population
M_POP	male population
TOT_SC	Scheduled Caste population
TOT_ST	Scheduled Tribe population
TOT_LIT	literate population