Preaching to the Choir: A Problem of Participatory Interventions
Codebook

Rebecca Hanson* Dorothy Kronick† Tara Slough‡

February 5, 2024

This codebook describes the processed datasets included in this replication package. All datasets described here are located in the `data` folder in the CodeOcean capsule or on Dataverse.

comm_policing_lit.Rdata

`comm_policing_lit.Rdata` (located in `data/`) contains three dataframes necessary to construct Figure A1, as detailed in Tables C1-C3.

motivation.Rdata

`motivation.Rdata` (located in `data/`) contains two dataframes necessary to construct Figure 2, as detailed in Tables C4-C5.

CuadranteTreatment.csv

`motivation.Rdata` (located in `data/`) contains one dataset necessary to construct Table A1, as detailed in Table C6. Note that the excel file `covariate_labs.xlsx` contains labels for these census variables that are used to construct Table A1. They do not constitute data.

survey_data.Rdata

`survey_data.Rdata` (located in `data/`) contains two dataframes that are used to conduct analysis of the (panel) survey data to construct most of the tables and figures in the manuscript and appendix, as detailed in Tables C7-C8. Note that the csv file `lasso_var_key.csv` contains labels used in the construction of the plot of the LASSO-selected coefficients (Figure A9). They do not constitute data.

mk_extract.Rdata

`mk_extract.Rdata` (located in `data/`) contains six dataframes, one for each of the Metaketa-IV experiments. They all follow the same structure, as described in Table C9.

lapop_sset.Rdata

`lapop_sset.Rdata` (located in `data/`) contains a dataframe used to conduct the cross-national analysis of LAPOP survey data, as detailed in Table C10. Note further that the dataframe `bootstrap_ests_lapop` (also located in `data/`) contain the output of bootstrapping exercise (500 observations of each quantity of interest) that are used to reduce computational time. Lines 100-115 of `05_lapop.R` detail the construction of these dataframes.

*r.hanson@ufl.edu
†kronick@berkeley.edu
‡tara.slough@nyu.edu

mtgs_gender.Rdata

mtgs_gender.Rdata (located in data/) contains a dataframe used to construct Figure A13, as described in Table C11. It is constructed from the qualitative meeting notes.

sentiment.Rdata

sentiment.Rdata (located in data/) contains a dataframe used to construct Figure A14, as detailed in Table C12.

Variable	Description	Complete obs.	Minimum	Maximum
year	Year of publication	52	1971	2022
gs	Articles indexed by Google Scholar (in 1000s)	52	0.01	4.39

Table C1: Variables in the dataframe gs.

Variable	Description	Complete obs.	Minimum	Maximum
year	Year of publication	219	1971	2022
ngrams	Community Policing' bigrams	219	0	0
ngrams_permillion	ngrams rescaled by a factor of 1,000,000	219	0	0.71

Table C2: Variables in the dataframe ngrams.

Variable	Description	Complete obs.	Minimum	Maximum
year	Year of study publication	40	1973	2021
count	Number of studies/evaluations published	40	1	20

Table C3: Variables in the dataframe ns.

Variable	Description	Complete obs.	Minimum	Maximum
`date`	Month	576	1998-01-01	2021-12-01
`group`	Geographic level of aggregation (Colombia or Antioquia)	482	28.51	172.34
`rate.12`	Number of homicides per 100k population over past 12 months in geographic unit	576		
`data`	Unit for homicide rate (Homicides per 100k population)	576		

Table C4: Variables in the dataframe `monthly_homicide`.

Variable	Description	Complete obs.	Minimum	Maximum
`year`	Year of LAPOP survey	24	2004	2018
`Sample`	Level of cross-sectional aggregation: either Colombia or Antioquia	24		
`trust`	Mean trust in police in wave in given cross-section	24	3.60	4.65

Table C5: Variables in the dataframe `lb_trust`.

Variable	Description	Complete obs.	Minimum	Maximum
cuadrante	Police beat administrative code	347		
estacion	Station that oversees the police beat	347	1	12
population	Population	347	119	19900
houses	Households	347	32	5737
per_hog	People per household	347	2.13	5.31
viv_arriendo	Share of rented homes	347	0	0.32
cuartos_dor	Avg. bedrooms per household	347	1	4.29
hacinamiento	Share of overcrowding households	347	0	0.42
ener_ele	Share of households cooking with electric energy	347	0	0.34
gas_pip	Share of households cooking with piped gas	347	0	1.01
vdom_bas	Share of households with fridge, TV, or sound system	347	0	0.37
comp	Share of households with computer	347	0	0.70
moto	Share of households with motorcycle	347	0	0.76
telefono	Share of households with landline	347	0.43	1
gas	Share of households with pipeline gas	347	0	0.97
ing_04	Adequate household income between 0 and 400,000 pesos	347	0	0.28
ing_410	Adequate household income between 400,000 and 1 mln pesos	347	0	0.48
ing_45	Adequate household income over 4.5mln pesos	347	0	0.11
empleados_hog	Number of employed people per household	347	0.29	1.79
desempleados_hog	Number of unemployed people per household	347	0.01	0.34
jefeh_obre	(Unclear from source documentation; omitted in analysis.)	347	0	0.07
pensionado_hog	Number of retirees per household	347	0	0.10
per_ext	Share of households with family member living abroad	347	0	0.19
per_hom	Share of males per household	347	0.10	0.61
per_muj	Share of females per household	347	0.11	0.64
per_menor	Share of under-aged per household	347	0.07	0.42
per_ter_edad	Share of seniors per household	347	0.02	0.80
jefeh_sexo	Share of households with male head	347	0.15	0.77

jefeh_solo_h	Share of households with single-male parent	347	0.02	0.43
jefeh_solo_m	Share of households with single-female parent	347	0.06	0.45
sinhijos	Share of households without children living at home	347	0.04	0.73
hijosh_edad_18	Share of households with university students	347	1.87	3.98
colombiano	Household head born in Colombia	347	0.21	1
medellin	Household head born in Medell\'in	347	0.09	0.64
chicago	Part of second study	347	0	9
logarriendo	Log of monthly rent	347	11	14.90
homicide	Homicide rate	347	0	1440.58
theft	Theft rate	347	0	508.93
laghomicide	Lagged homicide rate	347	0	352.94
lagtheft	Lagged theft rate	347	0	3976
chicagotreat	Second study treatment condition	347	0	1
chicagoctrl	Second study control condition	347	0	1
est1	Indicator for beat in station 1	347	0	1
est2	Indicator for beat in station 2	347	0	1
est3	Indicator for beat in station 3	347	0	1
est4	Indicator for beat in station 4	347	0	1
est5	Indicator for beat in station 5	347	0	1
est6	Indicator for beat in station 6	347	0	1
est7	Indicator for beat in station 7	347	0	1
est8	Indicator for beat in station 8	347	0	1
est9	Indicator for beat in station 9	347	0	1
est10	Indicator for beat in station 10	347	0	1
est11	Indicator for beat in station 11	347	0	1
blockids	Block indicator used in random assignment	347	1	87
treatment	Treatment assignment. Note that 1 = meeting and flyers, 2 = pure control, 3 = meetings only, and 4 = flyers only	347	1	4

Table C6: Variables in the file CuadranteTreatment.csv.

Variable	Description	Complete obs.	Minimum	Maximum
REGISTRO_BD	Respondent indicator for merging baseline and endline.	2434		
CUADRANTE	Police beat/cuadrante	3644	1	347
nro_cuadra	Police beat administrative code	3644		
block_ID	Block used for randomization	3644	1	87
Z_common	Indicator for meetings treatment	3644	0	1
Z_alt	Indicator for flyers treatment	3644	0	1
panel	Indicator for presence in both waves of survey	3644	0	1
any_exposure	Indicator for (self-reported) exposure to community-policing meetings in last 12 months	3644	0	1
attended	Indicator for (self-reported) attendance at community-policing meeting in last 12 months	3644	0	1
category	Categorical classification of exposure to community policing meetings	3644		
trust_police	Level of trust in police at endline (4 point scale)	3597	1	4
trust_po_bl	Level of trust in police at baseline (4 point scale)	2420	1	4
trust_bl_imp	Level of trust in police at baseline (4 point scale) with 0 imputed for missing observations	3644	0	4
trust_bl_miss	Indicator for missing baseline trust in police	3644	0	1
trust_beat_po	Level of trust in beat police at endline (4 point scale)	3476	1	4
po_index	Prespecified police quality index at endline(see pre-analysis plan for components). Standardized so that the mean is zero and standard deviation is 1 among pure control respondents.	3641	-2.34	2.26
po_index_bl	Prespecified police quality index from baseline survey (note that fewer components were measured at baseline) . Standardized so that the mean is zero and standard deviation is 1 among pure control respondents.	2431	-2.56	2.61
po_index_red	Prespecified police quality index from endline survey including only components measured at baseline . Standardized so that the mean is zero and standard deviation is 1 among pure control respondents.	3634	-1.99	1.90

po_index_bl_imp	Baseline police quality index with 0's imputed for missing observations.	3644	-2.56	2.61
po_index_bl_miss	Indicator for missing baseline police quality index	3644	0	1
po_beat_index	Prespecified police quality index for beat police at endline. Standardized so that the mean is zero and standard deviation is 1 among pure control respondents.	3610	-2.47	2.14
fisc_index	Prespecified prosecutor quality index at endline(see pre-analysis plan for components). Standardized so that the mean is zero and standard deviation is 1 among pure control respondents.	3612	-2.17	1.63
fisc_index_bl_imp	Prespecified prosecutor quality index from baseline survey (note that fewer components were measured at baseline) with 0's imputed for missing observations. Standardized so that the mean is zero and standard deviation is 1 among pure control respondents.	3644	-2.38	1.87
fisc_index_bl_miss	Indicator for missing baseline prosecutor quality index	3644	0	1
conviv_index	Convivencia index at endline. Standardized so that the mean is zero and standard deviation is 1 among pure control respondents.	3643	-2.36	1.94
conviv_index_bl	Convivencia index from baseline survey . Standardized so that the mean is zero and standard deviation is 1 among pure control respondents.	2434	-2.45	2.17
conviv_index_bl_imp	Convivencia index at baseline with 0's imputed for missing observations.	3644	-2.45	2.17
conviv_index_bl_miss	Indicator for missing baseline convivencia index	3644	0	1
sec_index	Prespecified security perceptions index at endline (see pre-analysis plan for components). Standardized so that the mean is zero and standard deviation is 1 among pure control respondents.	3644	-1.53	1.34

sec_index_bl	Prespecified security perceptions index from baseline survey . Standardized so that the mean is zero and standard deviation is 1 among pure control respondents.	2434	-1.61	1.38
sec_index_bl_imp	Security perceptions index with 0's imputed for missing baseline observations.	3644	-1.61	1.38
sec_index_bl_miss	Indicator for missing baseline security perceptions index	3644	0	1
act_appropriate_po	Response to "the police act upon citizen comments and com- plaints about security in my community" on a 1 to 5 scale where 5 = strongly agree.	3612	1	5
act_seriously_po	Response to "the police take cases seriously and investigate them." on a 1 to 5 scale where 5 = strongly agree.	3575	1	5
equal_service_po	Response to "the police provide the same quality of service to all citizens." on a 1 to 5 scale where 5 = strongly agree.	3575	1	5
capacity_investigate_po	Response to "the police have the capacity to investigate crimes and gather evidence effectively." on a 1 to 5 scale where 5 = strongly agree.	3583	1	5
capacity_respond_po	Response to "the police have the capacity to respond to incidents of crime in a timely manner." on a 1 to 5 scale where 5 = strongly agree.	3594	1	5
act_appropriate_po_bl	Baseline response to "the police act upon citizen comments and complaints about security in my community" on a 1 to 5 scale where 5 = strongly agree.	2406	1	5
act_seriously_po_bl	Baseline response to "the police take cases seriously and investigate them." on a 1 to 5 scale where 5 = strongly agree.	2359	1	5
equal_service_po_bl	Baseline response to "the police provide the same quality of service to all citizens." on a 1 to 5 scale where 5 = strongly agree.	2369	1	5

capacity_investigate_po_bl	Response to "the police have the capacity to investigate crimes and gather evidence effectively." on a 1 to 5 scale where 5 = strongly agree.	2366	1	5
capacity_respond_po_bl	Baseline response to "the police have the capacity to respond to incidents of crime in a timely manner." on a 1 to 5 scale where 5 = strongly agree.	2383	1	5
age	Respondent age (reported on baseline survey)	2430	15	90
migrant	Indicator for internal migrant to Medellín	2434	0	1

Table C7: Variables in the dataframe outcomes.

Variable	Description	Complete obs.	Minimum	Maximum
REGISTRO_BD	Respondent indicator (for merging baseline and endline)	5205	42	13976
CUADRANTE	Police beat/cuadrante	5205	1	347
nro_cuadra	Police beat administrative code	5205		
block_ID	Block used for random assignment.	5205	1	87
Z_common	Indicator for meetings treatment	5205	0	1
Z_alt	Indicator for flyers treatment	5205	0	1
in_endline	Indicator for whether respondent was recontacted at endline	5205	0	1
past_meeting	Indicator for reported having heard of a (past) community policing meeting at baseline. 1 indicates 'yes.'	5205	0	1
attended_past	Indicator for reported attending a (past) community policing meeting prior to baseline survey. 1 indicates 'yes'	5205	0	1
received_invite	Indicator for reported receiving an invite to a community-police meeting in the endline survey (a measure of compliance). 1 indicates 'yes'	2425	0	1
heard_of	Indicator for reported having heard of community-police meetings in the endline survey. 1 indicates 'yes'	1804	0	1
attendedv	Indicator for reported attending a community-police meeting in last 12 months in endline survey. 1 indicates 'yes'	2425	0	1
any_exposure	Indicator for having heard of (incl. receiving invite) or attended a community-police meeting in the endline survey	2434	0	1
male	Indicator for male respondent.	5205	0	1
estrato	Strata of dwelling (1 to 6 scale)	5205	1	6
estrato_servicios	Strata for of dwelling for public services (1 to 6 scale)	5205	1	6

education	Education level. Note that these categories are not ranked (e.g., 1 does imply lower education levels than 2). All analyses treat this variable as a factor.	5205	1	10
migrant	Indicator for internal migrant to Medellín. 1 indicates 'yes'	5205	0	1
decade	Age for respondent in decade bins (e.g., a respondent in their 20s would be coded as a 2).	5205	1	9
race	Self-identified race	5205		
COMUNA_LOCALIDAD	Comuna in which beat is located.	5205		
pleb	Self-reported vote in peace plebiscite.	5205		
not_partisan	Reported no partisan affiliation.	5205	0	1
centro_democratico	Reported support for the Centro Democratico party.	5205	0	1
farc	Reported support for Communes (the FARC-successor party)	5205	0	1
would_vote	Reported that they were "likely" or "very likely" to vote in (upcoming) presidential election if held today	5205	0	1
pres_vote_intention	Self-reported vote intention for 2018 presidential election (survey was conducted 3-6 months before election.	5205		
bl_trust	Baseline trust in police on 1 to 4 scale where 1 is low and 4 is high.	5171	1	4

Table C8: Variables in the dataframe survey.

Variable	Description	Complete obs.	Minimum	Maximum
clu	Cluster indicator	varies		
ZZ	Assignment to community-policing (harmonized) treatment	varies	0	1
mtg_attend	Indicator for (self-reported) meeting attendance	varies	0	1
satis_trust	Endline measure of trust in police	varies	varies	varies
satis_trust_bl	Baseline measure of trust in police	varies	varies	varies
in_baseline	Indicator for presence in baseline sample	varies	0	1

Table C9: Each of the data frames in the file mk_extract, including bra, col, lib, pak, phl, and uga contain a subset of these variables. Note that the Philippines site did not conduct a panel survey so does not have baseline variables. The Liberia survey used a repeated-cross section design, so the measure of baseline trust is a community-level average and there is no indicator for membership in baseline sample.

Variable	Description	Complete obs.	Minimum	Maximum
pais	Numeric country indicator (see LAPOP documentation or .dta file).	231,354	1	41
year	Year of survey.	231,354	1	41
wave	LAPOP wave of survey (redundant with year).	231,354	1	41
mtg	Response to survey question "Have you attended a town meeting, city council meeting or other meeting in the past 12 months?" 1 = yes and 0 = no	210,818	0	1
comm_assoc	Response to survey question "[Do you attend] meetings of a community improvement committee or association?" 1 = Any of the following responses: {once a week, once or twice a month, once or twice a year}, 0 = Never	231,354	1	41
mun_trust	Response to the survey question: "To what extent do you trust the local or municipal government?" (7-point scale rescaled to 0-1).	215,420	0	1
comm_trust	Response to the survey question: "And speaking of the people from around here, would you say that people in this community are very trustworthy, somewhat trustworthy, not very trustworthy or untrustworthy...?" (4-point scale flipped so that 4 is most trusting, and rescaled to 0-1).	224,412	1	41
ID	Country-wave indicator.	231,354		
unique_id	Respondent indicator	231,354		

Table C10: Variables in the dataframe lapop_sset.

Variable	Description	Complete obs.	Minimum	Maximum
NRO_CUADRA	Police beat administrative code (treatment beats only).	173		
mean_bl_trust	Average of meeting sentiment (Z-scores of iSOL) measure.	519	1	3
wm_share	Coded as (# female-presenting attendees - # male-presenting attendees)/total attendees. NA's mean that (1) meeting was cancelled or (2) gender of attendees was not reported in meeting notes.	519	-1	1

Table C11: Variables in the dataframe attend.

Variable	Description	Complete obs.	Minimum	Maximum
NRO_CUADRA	Police beat administrative code (treatment beats only)	173		
mean_bl_trust	Mean trust in police in baseline survey (on a 4-point scale)	173	1.8	3.2
mean_sentiment	Average of iSOL z-score and NRC emotion lexicon z-score for meeting notes.	118	-1.67	1.35
n_meetings	Number of community-police meetings held in beat	173	0	3

Table C12: Variables in the dataframe sentiment.