```
log using "20_data_process_log.txt", text replace
//// RCT Survey: Data Processing
//// Lucas Reddinger <jlr@lucasreddinger.com>, David Levine, Gary Charness
//// 2021 April 24

cls
clear
set more off

cd "E:\lucas\projects_private_git\academia\projects\vaccine\rct"

**********************
**** Setup Prolific data
**********************

** RCT 1 FINAL PROLIFIC SURVEY DATA
import delimited "data\data_prolific_survey_rct1_final.csv", varnames(1) rowrange(4)
destring _all, replace

rename participant_id prolific_pid

// We should not have any duplicates.
sort prolific_pid session_id
quietly by prolific_pid (session_id):  gen dup_pid = cond(_N==1,0,_n)
tab dup_pid

keep prolific_pid status age countryofbirth currentcountryofresidence ///
    employmentstatus firstlanguage nationality sex studentstatus

save "data\data_prolific_survey_rct1_final.dta", replace
clear

** RCT 2 RECENT PROLIFIC SURVEY DATA
import delimited "data\data_prolific_survey_rct2_2021-04-23_1134.csv", varnames(1) rowrange(4)
destring _all, replace

rename participant_id prolific_pid

// We should not have any duplicates.
sort prolific_pid session_id
quietly by prolific_pid (session_id):  gen dup_pid = cond(_N==1,0,_n)
tab dup_pid

keep prolific_pid status age countryofbirth currentcountryofresidence ///
    employmentstatus firstlanguage nationality sex studentstatus

save "data\data_prolific_survey_rct2_2021-04-23_1134.dta", replace
clear

*********************************
**** Import Qualtrics survey data
*********************************

import delimited "data\data_qualtrics_survey_2021-04-23_0939.csv", varnames(1) rowrange(4)
destring _all, replace

*********************
**** SHUFFLE OUTCOMES
*********************
// Obviously this will be removed when we are ready to analyze the data

set seed 578588
gen double shuffle1 = runiform()
gen double shuffle2 = runiform()

sort shuffle1
gen long which = _n
sort shuffle2
gen int manip_chk_new = manip_chk[which], after(manip_chk)
gen int vax_likely_new = vax_likely[which], after(vax_likely)
gen int vax_delay_new = vax_delay[which], after(vax_delay)
gen int vax_child_new = vax_child[which], after(vax_child)

list which vax_likely* in 1/30

drop manip_chk vax_likely vax_delay vax_child shuffle* which
rename manip_chk_new manip_chk
rename vax_likely_new vax_likely
```

```
rename vax_delay_new vax_delay
rename vax_child_new vax_child

*************************
**** Qualtrics data basics
*************************

**** Drop variables that have no information for us
drop startdate enddate responseid status progress recipient*
drop externalreference distributionchannel

**** These click data are not very interesting to us
drop *_firstclick *_lastclick *_clickcount

**** Drop some other junk
drop userlanguage religious*

**** Convert date-time to stata clock
gen double recordedclock = clock(recordeddate, "YMD hms"), after(recordeddate)
format recordedclock %tc
drop recordeddate

**** Recode race to black
gen d_black = 0, after(d_race)
replace d_black = 1 if d_race==2 | d_race==12
replace d_black = . if d_race>=.
drop d_race

**** Recode natural indicators
replace d_latinx = 0 if d_latinx == 2
replace d_parent = 0 if d_parent == 2

**** Rename d_smalltown_rural
rename d_s d_smalltown_rural

************************
**** Merge Prolific data
************************

** RCT 1 FINAL PROLIFIC SURVEY DATA
merge m:1 prolific_pid using "data\data_prolific_survey_rct1_final.dta", update
drop _merge

** RCT 2 RECENT PROLIFIC SURVEY DATA
merge m:1 prolific_pid using "data\data_prolific_survey_rct2_2021-04-23_1134.dta", update
drop _merge

*************************
**** Label variable values
*************************

#delimit ;

label variable manip_chk "How many people have died?";

label values manip_chk manip_chk_label;

label define manip_chk_label
        2 "Over 200,000"
        3 "Over 300,000"
        4 "Over 400,000"
        5 "Over 500,000"
        6 "Over 600,000";

label variable vax_likely "How likely get vax described above?";

label values vax_likely vax_likely_label;

label define vax_likely_label
        1 "Highly likely"
        2 "Somewhat likely"
        3 "Slightly likely"
        4 "Neither likely nor unlikely"
        5 "Slightly unlikely"
        6 "Somewhat unlikely"
        7 "Highly unlikely"
        8 "DK / prefer not to say";

```

```
label variable vax_delay "When would get vax described above?";

label values vax_delay vax_delay_label;

label define vax_delay_label
        1 "Already have taken a vaccine dose"
        2 "As soon as possible"
        3 "When more people have been vaccinated"
        4 "Unsure when I will take the vaccine"
        5 "Unlikely to ever take the vaccine"
        6 "Definitely will not take the vaccine"
        7 "DK / prefer not to say";

label variable covid_hist "Your experience w COVID-19 virus?";

label values covid_hist covid_hist_label;

label define covid_hist_label
        1 "Positive, overnight hospital stay"
        2 "Positive, no overnight hospital stay"
        3 "Strongly suspect was positive, not tested"
        4 "Strongly suspect was positive, tested negative"
        5 "Believe didn't have"
        6 "DK / prefer not to say";

label variable vax_child "How likely get child vax as above?";

label values vax_child vax_likely_label;

label variable d_reside "Residence";

label values d_reside d_reside_label;

label define d_reside_label
        1 "US"
        2 "Other";

label variable d_age "Age";

label values d_age d_age_label;

label define d_age_label
        1 "Under 18"
        2 "18 - 24"
        3 "25 - 34"
        4 "35 - 44"
        5 "45 - 54"
        6 "55 - 64"
        7 "65 - 74"
        8 "75 - 84"
        9 "85 or older";

label variable d_parent "Children";

label values d_parent d_parent_label;

label define d_parent_label
        0 "No"
        1 "Yes";

label variable d_black "Black?";

label values d_black d_black_label;

label define d_black_label
        0 "No"
        1 "Yes";

label variable d_latinx "Hispanic or Latinx?";

label values d_latinx d_latinx_label;

label define d_latinx_label
        0 "No"
        1 "Yes";

label variable d_gender "Gender?";

```

```
235  label values d_gender d_gender_label;
236
237  label define d_gender_label
238          1 "Male"
239          2 "Female"
240          3 "Other";
241
242  label variable d_vote "Vote in 2020?";
243
244  label values d_vote d_vote_label;
245
246  label define d_vote_label
247          1 "Biden"
248          2 "Trump"
249          3 "Other/Didn't vote/DK";
250
251  label variable d_pols "Political scale?";
252
253  label values d_pols d_pols_label;
254
255  label define d_pols_label
256          1 "Extremely lib"
257          2 "Liberal"
258          3 "Slightly lib"
259          4 "Moderate"
260          5 "Slightly cons"
261          6 "Conservative"
262          7 "Extremely cons"
263          8 "Don't know";
264
265  label variable d_relig "Religious preference?";
266
267  label values d_relig d_relig_label;
268
269  label define d_relig_label
270          1 "Pr - Evangelical"
271          2 "Pr - Mainline"
272          3 "Catholic"
273          4 "Mormon"
274          5 "Other Chr"
275          6 "Jewish"
276          7 "Muslim"
277          8 "Eastern"
278          9 "None"
279          10 "Other / DK";
280
281  label variable d_relobs "How often observe religion?";
282
283  label values d_relobs d_relobs_label;
284
285  label define d_relobs_label
286          1 "Sev/wk"
287          2 "Once/wk"
288          3 "Sev/mo"
289          4 "Once/mo"
290          5 "Sev/yr"
291          6 "Once/yr"
292          7 "LT once/yr"
293          8 "Never"
294          9 "DK / NA";
295
296  label variable d_popdens "Small town or rural?";
297
298  label values d_popdens d_popdens_label;
299
300  label define d_popdens_label
301          1 "City or suburb"
302          2 "Large town"
303          3 "Small town"
304          4 "Rural area";
305
306  #delimit cr
307
308  **************
309  **** Save data
310  **************
311
312  save "data\data_shuffled_2021-04-23_0939.dta", replace
```

```
313
314   *************
315   **** Codebook
316   *************
317
318   codebook
319
320   ***************
321   **** Close log
322   ***************
323
324   log close
325
```