Report on Results for Study 1

Jake Bowers

17 March 2022

0.1 The Design

This file reruns our principle study from `results_study1_first.Rmd` utilizing *last in time* responses. You'll find that the results are very similar to our principal analyses. They are included as a robustness check.

We created a research design where pairs of survey respondents were created so as to be nearly identical on ideology. These pairs were also created so that the overall research design compared favorably with a pair-randomized experiment and otherwise produced extremely homogenous comparisons in terms of age, education, gender, race, partisanship, family income, and missingness on family income.

```
library(here)
```

```
## here() starts at /Users/kwilso14/repo/brown/covid/COVID-VaccinesSocialNorms-private
```

```
source(here("src", "R", "000_constants_and_utils.R"))
source(here("src/R", "001_rmd_setup.R"))
```

```
## Loading required package: survival
```

```
## Loading required package: SparseM
```

```
##
## Attaching package: 'SparseM'
```

```
## The following object is masked from 'package:base':
##
##      backsolve
```

```
library(kableExtra)
library(tidyverse)
library(estimatr)
library(coin)
library(sensemakr)
library(sensitivitymult)

load(file = here(MATCHES_DIR, "dat_plus_matches_study1_last.rda"), verbose = TRUE)
```

```
Loading objects:
  dat3
  tab
  pair_diffs
  xbres_vars
  xbres
load(file = here(MATCHES_DIR, "outcome_analysis_study1_last.rda"))
```

Our design consists of 786 respondents from surveys 5–8 placed into pairs that were required to be the same on partisanship and survey. Is this a good research design for our question? Our main goal is to break the relationship between perceptions and potential confounders within pair — that is, we aim to create pairs such that the person who, say, perceives more of their family wearing masks is no more or less likely to be the older member of the pair (or the democrat in the pair, or the more well-educated member of the pair, etc..). Within pair comparisons of perceptions in such a design then cannot be said to be confounded by the variables that we balanced.

We present a few pieces of information about our design here.

First, we show balance on average. The following table shows that the pairs have nearly the same mean values on the key covariates that we aim to balance.

```
kable(tab[-nrow(tab), ])
```

	Mean 1	Mean 2	Diffs
ideo5	3.0687	3.0102	0.0585
dem_rep_oth	2.0534	2.0534	0.0000
agegood	48.1069	48.4860	-0.3791
female	0.5954	0.5954	0.0000
race_new	1.1756	1.1756	0.0000
faminc_new_Imp	6.4577	6.1830	0.2747
educ	3.7176	3.6616	0.0560
trust_in_science	3.9186	3.9198	-0.0013
covid_know	0.7693	0.7693	0.0000
covid_subj_know	3.6477	3.6477	0.0000
relig_scale	0.3303	0.3432	-0.0129

What about within-pair differences? The following table summarizes the differences within pairs.

```
pd_means <- sapply(pair_diffs, mean, na.rm = TRUE)
pd_quants <- sapply(pair_diffs, function(x) {
  quantile(abs(x), probs = c(0, .1, .25, .5, .75, .9, 1), na.rm = TRUE)
})

pd_tab <- rbind(pd_quants, pd_means)

kableExtra::kable(pd_tab[, -c(1, 9)], digits = 3)
```

	ideo	educ	age	female	race_new	dem_rep_oth	faminc	trust_in_science	covid_know	covid_know_NA	covi
0%	0.000	0.000	0.000	0	0	0	0.000	0.000	0.000	0	
10%	0.000	0.000	1.000	0	0	0	0.000	0.000	0.000	0	
25%	0.000	0.000	2.000	0	0	0	1.000	0.000	0.000	0	
50%	1.000	1.000	4.000	0	0	0	1.039	0.000	0.125	0	
75%	1.000	1.000	7.000	0	0	0	2.039	0.000	0.125	0	
90%	2.000	1.000	8.000	0	0	0	3.000	0.500	0.250	0	
100%	2.000	1.000	9.000	0	0	0	3.000	0.500	0.250	0	
pd_means	-0.053	0.071	-0.354	0	0	0	-0.051	0.011	-0.001	0	

As a final check, we compare our design to an established standard: a pair-randomized experiment. If we were to randomize which member of a pair perceived more versus less mask wearing in their social network, we would expect no systematic relationships between perceptions and these covariates within pairs — across many covariates, we would see some relationships, but if we were really to randomize, we would know exactly the distribution that these differences would take. This kind of firm knowledge of a standard for an unconfounded design suggests that we compare our non-randomized design to the equivalent randomized design in the process of arguing in favor of the design. Moreover @hansenbowers2008 showed that a single test can summarize the relationship between many covariates and a randomized intervention — thus allowing us to avoid the multiple comparisons problem of looking at many different covariate tests. That result for this design provides us with p=0.29 for the test of the hypothesis that the covariate-to-perception relationship is what we would observe in an experiment: that is, we have no strong argument against that null in this case. The subsequent table summarizes the variable-by-variable results. Again, we see that the person who is ranked as perceiving more positive health behaviors is no more or less likely to be older, etc.. That is, those covariates cannot strongly confound any comparisons we make within these pairs.

```
xbres$overall
```

```
        chisquare df p.value
matched     9.659  8  0.2898
```

```
xbres_vars <- data.frame(xbres$results[, c("rankperc=0", "rankperc=1", "adj.diff", "std.diff", "p"), "matched"])
xbres_vars$padj <- p.adjust(xbres_vars$p, method = "holm")
kable(arrange(xbres_vars, p), digits = 3)
```

	rankperc.0	rankperc.1	adj.diff	std.diff	p	padj
educ	3.654	3.725	0.071	0.049	0.057	0.624
agegood	48.473	48.120	-0.354	-0.020	0.182	1.000
trust__in__science	3.913	3.925	0.011	0.011	0.216	1.000
faminc__new.NATRUE	0.066	0.053	-0.013	-0.054	0.251	1.000
ideo5	3.066	3.013	-0.053	-0.039	0.319	1.000
relig__scale	0.333	0.340	0.007	0.038	0.500	1.000
faminc__new	6.344	6.293	-0.051	-0.016	0.590	1.000
covid__know	0.765	0.765	-0.001	-0.005	0.922	1.000
covid__subj__know	3.921	3.921	0.000	0.000	1.000	1.000
covid__know.NATRUE	0.224	0.224	0.000	0.000	1.000	1.000
covid__subj__know.NATRUE	0.776	0.776	0.000	0.000	1.000	1.000

0.2 Outcome Analysis

Our analysis of outcomes is quite simple: we regress reports of respondents' own intentions to be vaccinated, test, and wear masks on reports of those respondents' perception of such intentions in their social networks, conditional on pair. This amounts to comparing the own-intentions of the higher-perceiver within pair to the own-intentions of the lower-perceiver within pair and taking the average. We report HC2 standard errors.

I am just printing out the results here in raw form. First the scales of the variables to make interpretation easier:

```
somevars <- unique(c(
  all.vars(formula(lm_perc_fam)),
  all.vars(formula(lm_perc_neigh)),
  all.vars(formula(lm_perc_city)),
  all.vars(formula(lm_perc_state)),
  all.vars(formula(lm_mask_fam)),
  all.vars(formula(lm_test_fam))
))
```

```
kableExtra::kable(summary(dat3[, somevars]))
```

q1	q60__1a	q60__2a	q60__3a	q60__4a	q7	q100	q29	q25new
Min. :1.00	Min. :0.0	Min. :0.00	Min. :0.0	Min. :0.00	Min. :1.00	Min. :1.0	Min. :1.00	Min. :1.00
1st Qu.:3.00	1st Qu.:2.00	1st Qu.:2.00	1st Qu.:2.0	1st Qu.:2.00	1st Qu.:2.00	1st Qu.:3.0	1st Qu.:4.00	1st Qu.:2.00
Median :4.00	Median :3.0	Median :3.00	Median :3.0	Median :3.00	Median :4.00	Median :4.0	Median :5.00	Median :3.00
Mean :3.76	Mean :3.2	Mean :2.89	Mean :3.1	Mean :3.23	Mean :3.42	Mean :3.8	Mean :4.65	Mean :3.13
3rd Qu.:5.00	3rd Qu.:5.0	3rd Qu.:4.00	3rd Qu.:4.0	3rd Qu.:4.00	3rd Qu.:5.00	3rd Qu.:5.0	3rd Qu.:6.00	3rd Qu.:4.00
Max. :5.00	Max. :6.0	Max. :6.00	Max. :6.0	Max. :6.00	Max. :5.00	Max. :5.0	Max. :6.00	Max. :7.00
NA	NA	NA	NA	NA	NA	NA's :610	NA	NA

Perceptions of vaccination intentions of others strongly predicts own intentions — recall that we are very strongly holding constant ideology here.

```
lm_perc_fam
```

```
        Estimate Std. Error t value  Pr(>|t|) CI Lower CI Upper  DF
q60_1a    0.345    0.03051   11.31 7.654e-26    0.285    0.405 392
lm_perc_neigh
```

```
        Estimate Std. Error t value  Pr(>|t|) CI Lower CI Upper  DF
q60_2a   0.3054    0.03763   8.118 6.206e-15   0.2315   0.3794 392
lm_perc_city
```

```
        Estimate Std. Error t value  Pr(>|t|) CI Lower CI Upper  DF
q60_3a   0.2829    0.04104   6.894 2.181e-11   0.2022   0.3636 392
lm_perc_state
```

```
        Estimate Std. Error t value   Pr(>|t|) CI Lower CI Upper  DF
q60_4a   0.2245    0.04744   4.733 0.000003096   0.1312   0.3178 392
```

Notice the pattern: family predicts more strongly than neighborhood, and in turn neighborhood than city, and city than state. Could we see this by chance? In fact, we have fairly strong evidence against arguments that this pattern of effects increasing with more personal relationships are just from chance. For example, using a seemingly unrelated regression framework, we tested hypotheses of equality among the coefficients shown above

```
test_fam_vs_bigger
```

```
Linear hypothesis test (Chi^2 statistic of a Wald test)

Hypothesis:
```

```
fam_q60_1a_c - neigh_q60_2a_c = 0
fam_q60_1a_c - city_q60_3a_c = 0
fam_q60_1a_c - state_q60_4a_c = 0

Model 1: restricted model
Model 2: sursfit

  Res.Df Df Chisq Pr(>Chisq)
1   3139
2   3136  3    10       0.019 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
test_neigh_vs_bigger


Linear hypothesis test (Chi^2 statistic of a Wald test)

Hypothesis:
neigh_q60_2a_c - city_q60_3a_c = 0
neigh_q60_2a_c - state_q60_4a_c = 0

Model 1: restricted model
Model 2: sursfit

  Res.Df Df Chisq Pr(>Chisq)
1   3138
2   3136  2  3.68       0.16
test_city_vs_bigger


Linear hypothesis test (Chi^2 statistic of a Wald test)

Hypothesis:
city_q60_3a_c - state_q60_4a_c = 0

Model 1: restricted model
Model 2: sursfit

  Res.Df Df Chisq Pr(>Chisq)
1   3137
2   3136  1  1.74       0.19
```

The social perceptions of the family to intentions relationship is not restricted to vaccination intentions. It also works for mask wearing and testing.

```
lm_mask_fam

     Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper DF
q100   0.8694    0.05213   16.68 7.74e-29   0.7658    0.973 87
lm_test_fam

       Estimate Std. Error t value  Pr(>|t|) CI Lower CI Upper  DF
q25new   0.2231    0.05075   4.396 0.00001421   0.1233   0.3229 392
```

1 Sensitivity Analysis

This study was not randomized, but we are making statistical inferences as if it were — as if perceptions were randomly assigned within these pairs. We know that we have some biases here, but are limited in regards our data. Although we have adjusted for some of the largest drivers of perceptions, we have not managed to control them all. How large of an unobserved effect on perceptions have to be in order to over turn our results? We present two modes of sensitivity analysis here.

1.1 Hazlett and Cinelli

Our first approach builds on the approach in Cinelli and Hazlett in which they posit a linear regression model and imagine unobserved confounders. They propose a "Robustness Value" or a value of the influence of an unobserved confounder on **both** the outcome and treatment above which our substantive interpretation of our result would change. For example, we print out the results for the effect of perceptions of family below. The key piece for us is the `Robustness Value, q=1,alpha=.05`: this is the value above which we could no longer claim to detect effects (the other values are more liberal — values that would be required to make our relationships exactly zero in magnitude).

```r
sens_perc_fam <- sensemakr(
  estimate = lm_perc_fam$coef["q60_1a"],
  se = lm_perc_fam$std.error["q60_1a"], dof = lm_perc_fam$df
)
summary(sens_perc_fam)
```

```
Sensitivity Analysis to Unobserved Confounding

Model Formula: "Data provided manually"

Null hypothesis: q = 1 and reduce = TRUE
-- This means we are considering biases that reduce the absolute value of the current estimate.
-- The null hypothesis deemed problematic is H0:tau = 0

Unadjusted Estimates of 'D':
  Coef. estimate: 0.345
  Standard Error: 0.031
  t-value (H0:tau = 0): 11.31

Sensitivity Statistics:
  Partial R2 of treatment with outcome: 0.246
  Robustness Value, q = 1: 0.431
  Robustness Value, q = 1, alpha = 0.05: 0.373

Verbal interpretation of sensitivity statistics:

-- Partial R2 of the treatment with the outcome: an extreme confounder (orthogonal to the covariates) that explains 100% of the residual variance

-- Robustness Value, q = 1: unobserved confounders (orthogonal to the covariates) that explain more than 43.1% of the residual variance of both th

-- Robustness Value, q = 1, alpha = 0.05: unobserved confounders (orthogonal to the covariates) that explain more than 37.3% of the residual varia
sens_perc_fam
```

```
Sensitivity Analysis to Unobserved Confounding

Model Formula: "Data provided manually"

Null hypothesis: q = 1 and reduce = TRUE

Unadjusted Estimates of ' D ':
  Coef. estimate: 0.345
  Standard Error: 0.031
  t-value: 11.31

Sensitivity Statistics:
  Partial R2 of treatment with outcome: 0.246
  Robustness Value, q = 1 : 0.431
  Robustness Value, q = 1 alpha = 0.05 : 0.373

For more information, check summary.
```

So, we would need some covariates that are unrelated the variables included in the pairing that would explain more than `r round(100*sens_perc_fam$sensitivity_stats$rv_qa,1)`% of the variation in **both** the outcome **and** perceptions in order for our statistical tests to not-reject the null of no effects.

The other Robustness Values are displayed below.

All of the values are pretty large: what can we imagine that could predict perceptions and vaccination intention independently of all of the variables in the pairing as strongly as would be required below?

```r
get_rv_qa_value <- function(model) {
  res <- sensemakr(
    estimate = model$coef[1],
    se = lm_perc_fam$std.error[1], dof = lm_perc_fam$df
  )
  return(res$sensitivity_stats$rv_qa)
}

robustness_values <- sapply(ls(patt = "^lm_mask|^lm_perc|^lm_test"), function(lmname) {
  get_rv_qa_value(get(lmname))
})

sort(robustness_values, decreasing = TRUE)
```

```
  lm_mask_fam   lm_perc_fam lm_perc_neigh  lm_perc_city lm_perc_state    lm_test_fam
       0.7151        0.3733        0.3319        0.3071        0.2376         0.2359
```

1.2 Rosenbaum Style

Paul Rosenbaum developed an approach to sensitivity analysis that Hazlett and Cinelli build on. It involves positing an unobserved factor that, within pair, changes the probability of "selection into treatment" away from uniform. Our "treatment" variable here has 7 categories, so we change the analysis a bit here: the question becomes whether the person who perceives more within a pair (rank of perceptions = 1) also tends to be the person with more positive vaccination intentions. In this formulation, if two people do not differ in perceptions, then this pair adds nothing to the analysis and they are dropped. The table below shows the estimates and p-values from a regression of outcomes on within-pair rank in perception.

```r
dat3 <- dat3 %>%
  group_by(bm) %>%
  mutate(
    q60_1a_rank = rank(q60_1a) - 1,
    q60_2a_rank = rank(q60_2a) - 1,
    q60_3a_rank = rank(q60_3a) - 1,
    q60_4a_rank = rank(q60_4a) - 1,
    q100_rank = rank(q100) - 1,
    q25_new_rank = rank(q25new) - 1
  ) %>%
  ungroup()

## Notice that sometimes the higher perceiver has a 1 and sometimes a 6 (mostly higher numbers of course).
with(dat3, table(q60_1a_rank, q60_1a, exclude = c()))
```

```
           q60_1a
q60_1a_rank   0   1   2   3   4   5   6
        0    34 112  44  76  41  16   0
        0.5  10  42   2  22  10  46   8
        1     0  10  13  66  48 127  59
```

```r
lm_perc_fam_rank <- lm_robust(q1 ~ q60_1a_rank, fixed_effects = ~bm, data = dat3)
lm_perc_neigh_rank <- lm_robust(q1 ~ q60_2a_rank, fixed_effects = ~bm, data = dat3)
lm_perc_city_rank <- lm_robust(q1 ~ q60_3a_rank, fixed_effects = ~bm, data = dat3)
lm_perc_state_rank <- lm_robust(q1 ~ q60_4a_rank, fixed_effects = ~bm, data = dat3)
lm_mask_fam_rank <- lm_robust(q7 ~ q100_rank, fixed_effects = ~bm, data = dat3)
lm_test_fam_rank <- lm_robust(q29 ~ q25_new_rank, fixed_effects = ~bm, data = dat3)

## These are just to help us understand the results below:
## the pattern of estimates and statistical tests should be the same as when we use the full scale
lm_rank_res_lst <- lapply(ls(patt = "_rank$"), function(lmnm) {
  res <- tidy(get(lmnm))
  res$model <- lmnm
  return(res)
})
lm_rank_res <- bind_rows(lm_rank_res_lst) %>%
  select(term, estimate, p.value, conf.low, conf.high, model, outcome) %>%
  mutate(across(where(is.numeric), round, 3))
lm_rank_res
```

```
       term estimate p.value conf.low conf.high              model outcome
1    q100_rank    0.117   0.239   -0.078     0.312    lm_mask_fam_rank      q7
2  q60_3a_rank    0.512   0.000    0.345     0.680   lm_perc_city_rank      q1
3  q60_1a_rank    0.780   0.000    0.619     0.941    lm_perc_fam_rank      q1
4  q60_2a_rank    0.647   0.000    0.484     0.810  lm_perc_neigh_rank      q1
5  q60_4a_rank    0.407   0.000    0.224     0.591  lm_perc_state_rank      q1
6 q25_new_rank    0.578   0.000    0.345     0.810    lm_test_fam_rank     q29
```

```r
## here is the p-value with no bias
find_gamma_lim(g = 1, y = "q1", dat = dat3, z = "q60_1a_rank", return = "all")$pval
```

```
[1] 0
```

```r
lm_perc_fam_rank ## notice basically same
```

```
            Estimate Std. Error t value  Pr(>|t|) CI Lower CI Upper  DF
q60_1a_rank   0.7802    0.08177   9.541 1.536e-19   0.6194   0.9409 392
```

```r
## Now find the Gamma (or size of the selection effect) required to drive p>=.05
sens_perc_fam_G <- uniroot(f = find_gamma_lim, y = "q1", z = "q60_1a_rank", dat = dat3, lower = 1, upper = 20)
## Here is the Gamma value: the covariate has to cause the person to be 4 times more likely to be the higher perceiver within the same pair
sens_perc_fam_G$root
```

```
[1] 3.493
```

```r
## Showing the the pvalue is .0501 in this case
find_gamma_lim(g = sens_perc_fam_G$root, y = "q1", z = "q60_1a_rank", dat = dat3, return = "all")$pval
```

```
[1] 0.05099
```

```
## Now for the other variables
gamma_vals <- sapply(1:nrow(lm_rank_res), function(i) {
  message(lm_rank_res[i, "outcome"], " ", lm_rank_res[i, "term"])
  res_G <- try(uniroot(f = find_gamma_lim, y = lm_rank_res[i, "outcome"], z = lm_rank_res[i, "term"], dat = dat3, lower = 1, upper = 20))
  if (inherits(x = res_G, "try-error")) {
    res_G$root <- NA
  }
  return(res_G$root)
})
```

```
Error in uniroot(f = find_gamma_lim, y = lm_rank_res[i, "outcome"], z = lm_rank_res[i,  :
  f() values at end points not of opposite sign
```

```
Warning in res_G$root <- NA: Coercing LHS to a list
```

```
names(gamma_vals) <- lm_rank_res$term[1:nrow(lm_rank_res)]
sort(gamma_vals, decreasing = TRUE, na.last = TRUE)
```

```
 q60_1a_rank  q60_2a_rank  q60_3a_rank q25_new_rank  q60_4a_rank     q100_rank
       3.493        2.599        1.994        1.721        1.476            NA
```

Notice that the stronger the relationship, the larger the bias required to drive the p-value above .05. Strongest effect is family, for example.

This next is the Rosenbaum and Silber approach to interpreting sensitivity parameters. Notice that the basic gamma values assume a gigantic relationship with the outcome.

```
amps <- lapply(gamma_vals, function(theg) {
  if (is.na(theg)) {
    return(NA)
  }
  message(theg)
  my_amplify(theg, lambda = seq(theg + .001, 10))
})
names(amps) <- names(gamma_vals)
```

```
amps
```

```
$q100_rank
[1] NA

$q60_3a_rank
        delta lambda
 [1,] 2977.401  1.995
 [2,]    4.966  2.995
 [3,]    3.481  3.995
 [4,]    2.985  4.995
 [5,]    2.738  5.995
 [6,]    2.589  6.995
 [7,]    2.490  7.995
 [8,]    2.419  8.995
 [9,]    2.366  9.995

$q60_1a_rank
         delta lambda
[1,] 11204.217  3.494
[2,]    14.682  4.494
[3,]     9.091  5.494
[4,]     7.225  6.494
[5,]     6.292  7.494
[6,]     5.733  8.494
[7,]     5.359  9.494

$q60_2a_rank
        delta lambda
[1,] 5758.837    2.6
[2,]    8.350    3.6
[3,]    5.476    4.6
[4,]    4.517    5.6
[5,]    4.038    6.6
[6,]    3.750    7.6
[7,]    3.558    8.6
[8,]    3.421    9.6

$q60_4a_rank
        delta lambda
[1,] 1180.660  1.477
[2,]    2.654  2.477
[3,]    2.066  3.477
[4,]    1.869  4.477
```

```
[5,]    1.771  5.477
[6,]    1.712  6.477
[7,]    1.673  7.477
[8,]    1.645  8.477
[9,]    1.624  9.477


$q25_new_rank
        delta lambda
[1,] 1964.232  1.722
[2,]    3.682  2.722
[3,]    2.702  3.722
[4,]    2.375  4.722
[5,]    2.212  5.722
[6,]    2.114  6.722
[7,]    2.048  7.722
[8,]    2.002  8.722
[9,]    1.966  9.722
```

The p-value that we see for a given Γ (like `r gamma_vals[[1]]`) can be produced from an unobserved covariate that increases the treatment odds (within all pairs) by lambda and increases the odds of a positive pair difference in outcomes by delta.

1.3   Auxiliary analyses

What if, in every pair, the person who reports that she perceives more vaccination support among her family and friends, neighborhors, and people in her city and state does so because she *hopes* that they will be vaccinated — that is, she herself believes vaccination is good, so she hopes that those close to her will be vaccinated. This kind of wishful thinking or imagining a consensus almost certainly does not occur in *every* pair, but yet, we wonder in how many pairs we might see this phenomenon. Notice that saying "no one will vaccinate" for all levels and saying "everyone will vaccinate" at all levels both indicates a kind of wishful thinking (either believing that vaccination is bad and no one should do it, and believeing that it is good that everyone should do it.) How many people exhibit this kind of response pattern? Do people with these response patterns tend to also report either a stronglya negative or positive own intention? If few do, and this pattern does not strongly predict own intention, then we have less concern about wishful thinking driving our results.

Furthermore, we should note that this kind of false consensensus or reverse causality would not necessarily imply the pattern of results that we see: someone who is not actually reporting perceptions but hopes or rationalizations is likely to project those to all groups, and not first to family and friends (or most strongly to them) and then less strongly as the group becomes more distant and diffuse.

We see here that very few people show this response pattern, too few to overturn our results (considering the sensitivity analysis above). Also, notice that those having this pattern are not perfectly homogenous in their own intentions — those reporting that all will definitely be vaccinated at all levels are very likely to say that they will be vaccinated, and the converse is also true, but this is not a perfect relationship.

```
table(dat3$q60_1a, exclude = c())


  0   1   2   3   4   5   6
 44 164  59 164  99 189  67
table(dat3$q60_2a, exclude = c())


  0   1   2   3   4   5   6
 32 137 126 231 136 101  23
table(dat3$q60_3a, exclude = c())


  0   1   2   3   4   5   6
 16  97 119 259 169 108  18
table(dat3$q60_4a, exclude = c())


  0   1   2   3   4   5   6
 13  55  97 317 190 102  12
dat3 <- dat3 %>% mutate(
  perc_all_vax = as.numeric(q60_1a == 6 & q60_2a == 6 & q60_3a == 6 & q60_4a == 6),
  perc_none_vax = as.numeric(q60_1a == 0 & q60_2a == 0 & q60_3a == 0 & q60_4a == 0)
)
```

```
## Very few people have this response pattern:
with(dat3, table(perc_all_vax, exclude = c()))
```

```
perc_all_vax
  0   1
778   8
```
```
with(dat3, table(perc_none_vax, exclude = c()))
```

```
perc_none_vax
  0   1
777   9
```
```
with(dat3, table(perc_all_vax, q1, exclude = c()))
```

```
            q1
perc_all_vax   1   2   3   4   5
           0  89  48 151 167 323
           1   0   1   0   1   6
```
```
with(dat3, table(perc_none_vax, q1, exclude = c()))
```

```
             q1
perc_none_vax   1   2   3   4   5
            0  85  49 149 165 329
            1   4   0   2   3   0
```

2   References