Review_appendix

2023-07-12

Contents

Review 1

Comment 1

Comment Third, how can the authors differentiate the impact of intergroup contact from the impact of the development project around which the contact was organized? Treatment communities received large monetary contributions to be used for development projects that could potentially reduce the resource scarcity that is driving conflict (water access points and fencing to protect crops from grazing livestock). Maybe the reported findings have nothing to do with intergroup contact and are instead driven by a reduction in the triggers for conflict. The authors acknowledge this possibility on page 6, but only within the context of trying to understand diffusion mechanisms. While the experiment cannot be rerun, there may be analyses that could help address this (e.g., if the effects are larger in places with greater scarcity, this would suggest that the resource infusion is doing much of the work).

Response Great question. We might be able to answer this that the boreholes were largely constructed at the end of the project, and likely had little impact on water scarcity at the time of the survey. Participants might have had expectations the scarcity would be lifted, and that's hard to disentangle. Not sure what analysis we could do to disentangle. Maybe the fact that at endline, in Benue, pastoralists would not have benefited from the borehole as they were displaced?

- Few were aware of mediation

- Less likely to have direct short-term effects on attitudes
- Look at people who used the boreholes
- Maybe look at Benue pastoralists: is TR effect smaller there? (issue: low power)

Code

Few people participated in negotiations, relatively few were even aware of the ECPN development project or of Mercy Corps.

```
## mediation awareness/participation
table(rand.df$ecpn_group.negotiation, exclude=c())
```

```
##
##    0    1 <NA>
##   35   43 3013
```

```
table(rand.df$heard_MC_group.heard_MC, exclude=c())
```

```
##
##    0    1 <NA>
##  979  559 1553
```

```
table(rand.df$heard_MC_group.part_MC, exclude=c())
```

```
##
##    0    1 <NA>
##  475   79 2537
```

Bigger effects among people aware/used QIP, people aware/used borehole? Among TR group, does qip/borehole awareness and use predict improved outcomes?

```
##
##    0    1 <NA>
##  402  648 2041
```

```
##
##    0    1 <NA>
##  639  409 2043
```

Show the analysis

```
# all (hard to determine a trend across so many tests)
print(qipList)
```

```
## $`qip_group.borehole_use_end and attitude_cw`
##                                                      term   estimate    p.value
## 1                                             (Intercept) -0.1551652 0.25134306
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])]  0.4279773 0.09977711
##   statistic
## 1 -1.266758
```

```
## 2  1.953857
##
## $`qip_group.borehole_aware_end and attitude_cw`
##                                                       term    estimate   p.value
## 1                                              (Intercept) -0.09786287 0.5847451
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])]  0.17770567 0.4213694
##    statistic
## 1 -0.5791416
## 2  0.8543039
##
## $`qip_ben_end and attitude_cw`
##                                                       term   estimate   p.value
## 1                                              (Intercept)  0.1228893 0.2441145
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])] -0.2321103 0.1593870
##   statistic
## 1  1.383928
## 2 -1.730619
##
## $`qip_aware_end and attitude_cw`
##                                                       term   estimate    p.value
## 1                                              (Intercept) -0.2393563 0.10737179
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])]  0.4008580 0.05417202
##   statistic
## 1 -2.016055
## 2  2.545170
##
## $`qip_group.borehole_use_end and in_cw`
##                                                       term  estimate   p.value
## 1                                              (Intercept) 0.1252436 0.2319879
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])] 0.2252874 0.2411557
##   statistic
## 1  1.326710
## 2  1.304142
##
## $`qip_group.borehole_aware_end and in_cw`
##                                                       term  estimate   p.value
## 1                                              (Intercept) 0.1137316 0.4346263
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])] 0.1597918 0.4306371
##   statistic
## 1 0.8405028
## 2 0.8365395
##
## $`qip_ben_end and in_cw`
##                                                       term  estimate   p.value
## 1                                              (Intercept) 0.0667589 0.6761734
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])] 0.2412950 0.3300336
##   statistic
## 1 0.4525381
## 2 1.1096260
##
## $`qip_aware_end and in_cw`
##                                                       term     estimate   p.value
## 1                                              (Intercept) 0.0001419501 0.9991917
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])] 0.3532101220 0.1159091
```

```
##     statistic
## 1 0.001071637
## 2 1.921185223
##
## $`qip_group.borehole_use_end and contactOnly_cw`
##                                                  term   estimate   p.value
## 1                                         (Intercept) -0.1521069 0.2905232
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])]  0.4611912 0.1504413
##   statistic
## 1 -1.156818
## 2  1.654246
##
## $`qip_group.borehole_aware_end and contactOnly_cw`
##                                                  term   estimate   p.value
## 1                                         (Intercept) -0.1016586 0.5837842
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])]  0.2094608 0.3941083
##    statistic
## 1 -0.5806665
## 2  0.9082565
##
## $`qip_ben_end and contactOnly_cw`
##                                                  term    estimate   p.value
## 1                                         (Intercept) -0.05203494 0.6813960
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])]  0.12239958 0.4061062
##    statistic
## 1 -0.4446118
## 2  0.9288455
##
## $`qip_aware_end and contactOnly_cw`
##                                                  term   estimate   p.value
## 1                                         (Intercept) -0.2647398 0.1758762
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])]  0.4899805 0.1066462
##   statistic
## 1 -1.610426
## 2  1.987892
##
## $`qip_group.borehole_use_end and rMean`
##                                                  term   estimate    p.value
## 1                                         (Intercept) -0.2076885 0.11796895
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])]  0.6913695 0.01664858
##   statistic
## 1 -1.818278
## 2  3.318732
##
## $`qip_group.borehole_aware_end and rMean`
##                                                  term   estimate   p.value
## 1                                         (Intercept) -0.1827557 0.3721768
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])]  0.3945841 0.1180975
##    statistic
## 1 -0.9686963
## 2  1.7826027
##
## $`qip_ben_end and rMean`
##                                                  term   estimate   p.value
```

```
## 1                                           (Intercept)  0.1809871 0.2864994
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])] -0.2321093 0.2209224
##   statistic
## 1  1.244310
## 2 -1.452014
##
## $`qip_aware_end and rMean`
##                                                   term   estimate    p.value
## 1                                           (Intercept) -0.3143459 0.05854819
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])]  0.6320611 0.01050003
##   statistic
## 1 -2.535441
## 2  4.095751
##
## $`qip_group.borehole_use_end and end_exp`
##                                                   term    estimate   p.value
## 1                                           (Intercept) -0.04596355 0.8460713
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])] -0.12814460 0.7659214
##    statistic
## 1 -0.2025309
## 2 -0.3119195
##
## $`qip_group.borehole_aware_end and end_exp`
##                                                   term   estimate   p.value
## 1                                           (Intercept) -0.2521245 0.4307204
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])]  0.2472291 0.5751816
##    statistic
## 1 -0.8480947
## 2  0.5878861
##
## $`qip_ben_end and end_exp`
##                                                   term   estimate   p.value
## 1                                           (Intercept) -0.2724795 0.2545324
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])]  0.4291659 0.3871612
##    statistic
## 1 -1.3474947
## 2  0.9709019
##
## $`qip_aware_end and end_exp`
##                                                   term    estimate   p.value
## 1                                           (Intercept) -0.03574193 0.8906272
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])]  0.01111851 0.9833447
##    statistic
## 1 -0.1456248
## 2  0.0219986
##
## $`qip_group.borehole_use_end and pgp_donate_end`
##                                                   term    estimate
## 1                                           (Intercept)  0.98066515
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])] -0.02456554
##        p.value  statistic
## 1 7.281328e-11 92.7762275
## 2 4.064013e-01 -0.8944117
##
```

```
## $`qip_group.borehole_aware_end and pgp_donate_end`
##                                                    term    estimate
## 1                                           (Intercept)  0.99457660
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])] -0.03754193
##        p.value statistic
## 1 9.738331e-11  113.5298
## 2 6.362711e-02   -2.2035
##
## $`qip_ben_end and pgp_donate_end`
##                                                    term    estimate
## 1                                           (Intercept)  0.98264149
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])] -0.02439341
##        p.value  statistic
## 1 6.270997e-07 71.0561110
## 2 3.852131e-01 -0.9753259
##
## $`qip_aware_end and pgp_donate_end`
##                                                    term    estimate
## 1                                           (Intercept)  0.98968360
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])] -0.03624162
##        p.value statistic
## 1 9.563338e-08 64.744984
## 2 2.502487e-01 -1.309436
##
## $`qip_group.borehole_use_end and pgp_amount_end`
##                                                    term estimate      p.value
## 1                                           (Intercept) 291.2945 1.051305e-05
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])] -55.0763 3.372547e-01
##   statistic
## 1 13.101154
## 2 -1.045054
##
## $`qip_group.borehole_aware_end and pgp_amount_end`
##                                                    term  estimate      p.value
## 1                                           (Intercept) 290.83344 2.735736e-05
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])] -33.85793 3.758690e-01
##    statistic
## 1 12.2273019
## 2 -0.9459166
##
## $`qip_ben_end and pgp_amount_end`
##                                                    term   estimate     p.value
## 1                                           (Intercept) 272.444849 0.000296108
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])]  -3.518875 0.928148517
##     statistic
## 1 13.30701211
## 2 -0.09605319
##
## $`qip_aware_end and pgp_amount_end`
##                                                    term estimate      p.value
## 1                                           (Intercept) 299.9294 0.0002920901
## 2 ag.df[ag.df$treatment %in% 1, paste(ben_out_df[i, 1])] -51.2100 0.3427752088
##   statistic
## 1 10.345767
```

```
## 2 -1.053596
```

```
# coefficients for each outcome across "benefit" variables
## no meaningful trend, but most coefficients positive and pvals below 0.50, so somewhat suggestive tha
benefit_df
```

```
##                        coefs      pvals      tstat
## attitude          0.19360767 0.18367638  0.9056779
## in                0.24489606 0.27943385  1.2928731
## contactOnly       0.32075802 0.26432550  1.3698100
## rMean             0.37147635 0.09154214  1.9362679
## end_exp           0.13984223 0.67790222  0.3172168
## pgp_donate_end   -0.03068563 0.27637253 -1.3456686
## pgp_amount_end  -35.91577624 0.49601187 -0.7851551
## all                       NA 0.32418064  0.5272889
```

```
# coefficients for each "benefit variables" across outcomes
## no meaningful trend, but coef/tstat/pval could be affected by PGG being weird negative outcome
benefitVar_df
```

```
##                                  coefs     pvals      tstat
## qip_group.borehole_use_end   -7.6318833 0.2882286  0.8542273
## qip_group.borehole_aware_end -4.6723864 0.3398414  0.2600245
## qip_ben_end                  -0.4592325 0.4024246 -0.1778056
## qip_aware_end                -7.0512873 0.2662280  1.1727092
## all                                  NA 0.3241806  0.5272889
```

```
# coefficients for each "benefit variables" across outcomes, removing PGG
## no clear trend, seems slightly suggestive that benefiting weakly improved outcomes.
benefitVar_df_svy
```

```
##                                  coefs     pvals       tstat
## qip_group.borehole_use_end   0.33553615 0.2547888  1.58381134
## qip_group.borehole_aware_end 0.23775427 0.3878788  0.99391776
## qip_ben_end                  0.06572817 0.3007221 -0.03465197
## qip_aware_end                0.37744566 0.2541144  2.11439939
## all                                  NA 0.2993760  1.16436913
```

Unecessary Robustness check

```
# obvious: control group is all NA or 0 for borehole use/aware.
table(ag.df$qip_group.borehole_use_end>0, ag.df$treatment, exclude=c())
```

```
##
##          0  1
##   FALSE  1  0
##   TRUE   0 20
##   <NA>   9  0
```

```
# robustness checks
## make a version thats all thats all 0 if NA (NA really means they didnt benefit)
ag.df$qip_group.borehole_use_end <- ifelse(is.na(ag.df$qip_group.borehole_use_end), 0, ag.df$qip_group.b
ag.df$qip_group.borehole_aware_end <- ifelse(is.na(ag.df$qip_group.borehole_aware_end), 0, ag.df$qip_gro
ag.df$qip_ben_end <- ifelse(is.na(ag.df$qip_ben_end), 0, ag.df$qip_ben_end)
ag.df$qip_aware_end <- ifelse(is.na(ag.df$qip_aware_end), 0, ag.df$qip_aware_end)

# run above code with robustness check var, save lists with "_0" at end.
load(file="qip_list_0.Rda")
load(file="plot_list_0.Rda")
load(file="benefit_df.Rda")
load(file="benefitVar_df.Rda")
```

Benue pastoralists less likely to perceive benefit because (were displaced, not using boreholes). Are effects among Benue pastoralists different than other respondents?

No patterns.

chris: more here. (did more but could do even more)

```
# make ben past variable
ag.df$ben_past <- ifelse(ag.df$state %in% "ben" & ag.df$farm_past %in% "past", 1, 0)

# Benue pastoralists: are they less aware and do they use/benefit from projects less?
## yes, they are less aware and use the borehole less (mixed on other project, which was done earlier)
lm_robust(qip_group.borehole_use_end~ben_past,
                            clusters = psu, data=ag.df[ag.df$treatment %in% 1,])
```

```
##                Estimate Std. Error   t value     Pr(>|t|)   CI Lower   CI Upper
## (Intercept)  0.4632439 0.06076430  7.623618 5.587948e-05  0.3235646  0.6029232
## ben_past    -0.3406948 0.06563456 -5.190784 5.832725e-03 -0.5199676 -0.1614221
##                    DF
## (Intercept) 8.148559
## ben_past    4.174049
```

```
lm_robust(qip_group.borehole_aware_end~ben_past,
                            clusters = psu, data=ag.df[ag.df$treatment %in% 1,])
```

```
##                Estimate Std. Error   t value     Pr(>|t|)   CI Lower   CI Upper
## (Intercept)  0.7152896 0.08396237  8.519169 2.476459e-05  0.5222848  0.9082945
## ben_past    -0.4309759 0.09939244 -4.336103 1.118970e-02 -0.7024542 -0.1594976
##                    DF
## (Intercept) 8.148559
## ben_past    4.174049
```

```
lm_robust(qip_aware_end~ben_past,
                            clusters = psu, data=ag.df[ag.df$treatment %in% 1,])
```

```
##                Estimate Std. Error   t value     Pr(>|t|)   CI Lower   CI Upper
## (Intercept)  0.5762138 0.09732716  5.920380 0.0003292883  0.3524872 0.79994036
## ben_past    -0.2907236 0.11767188 -2.470629 0.0662888960 -0.6121299 0.03068276
##                    DF
## (Intercept) 8.148559
## ben_past    4.174049
```

```r
lm_robust(qip_ben_end~ben_past,
                            clusters = psu, data=ag.df[ag.df$treatment %in% 1,])
```

```
##                 Estimate Std. Error    t value    Pr(>|t|)   CI Lower  CI Upper
## (Intercept)  0.52276923 0.08578581  6.0938892 0.000270639  0.3255728 0.7199656
## ben_past    -0.06443589 0.10115950 -0.6369732 0.557407490 -0.3407407 0.2118689
##                   DF
## (Intercept) 8.148559
## ben_past    4.174049
```

```r
# benue pastoralist differences?
## is the TR effect weaker among Benue pastoralists?
## chris: is this reg correct?  Or should it just be all benue pastoralists?  Or should it only be TR Be
benPastList <- vector(mode="list", length=length(outcome_list_qip))
names(benPastList) <- paste(outcome_list_qip)
for(i in 1:length(outcome_list_qip))
{
  # thelm
  thelm <- lm_robust(ag.df[,outcome_list_qip[i]]~ben_past*treatment,
                     clusters = psu, data=ag.df)

  # tidy it, grab term, coef, pval
  lm_res <- tidy(thelm)[, c(1,2,5,4)]

  # save to list
  benPastList[[i]] <- lm_res
}

save(benPastList, file="benPast_list.Rda")
```

```r
print(benPastList)
```

```
## $attitude_cw
##                 term    estimate   p.value  statistic
## 1        (Intercept) -0.04413259 0.5909207 -0.5890474
## 2           ben_past -0.21101793 0.1646222 -2.6472167
## 3          treatment  0.10661302 0.2844967  1.1553974
## 4 ben_past:treatment -0.03173032 0.8260431 -0.2398319
##
## $in_cw
##                 term    estimate    p.value  statistic
## 1        (Intercept)  0.04896055 0.54884848  0.6606510
## 2           ben_past  0.03102023 0.92663677  0.1087160
## 3          treatment  0.19246330 0.05314168  2.3063954
## 4 ben_past:treatment -0.16686096 0.64004837 -0.5187857
##
## $contactOnly_cw
##                 term     estimate    p.value   statistic
## 1        (Intercept)  0.004148459 0.93281436  0.09036735
## 2           ben_past -0.559077279 0.08826057 -4.25210716
## 3          treatment  0.090914094 0.25088566  1.24756973
## 4 ben_past:treatment  0.234324708 0.31880309  1.19451968
```

```
##
## $rMean
##                 term    estimate   p.value  statistic
## 1        (Intercept)  0.09490267 0.4657685  0.8149186
## 2           ben_past -0.45780357 0.1136704 -3.5228563
## 3          treatment  0.05610774 0.6756433  0.4357864
## 4 ben_past:treatment  0.03012622 0.8672572  0.1820363
##
## $end_exp
##                 term   estimate     p.value statistic
## 1        (Intercept) -0.3802838 0.005114448 -6.123002
## 2           ben_past  0.8058674 0.130915463  3.164668
## 3          treatment  0.2239975 0.139017323  1.661355
## 4 ben_past:treatment -0.5074066 0.196886090 -1.657151
##
## $pgp_donate_end
##                 term    estimate      p.value  statistic
## 1        (Intercept)  0.93567402 8.732914e-05 19.7087029
## 2           ben_past  0.06432598 3.499128e-01  1.3549394
## 3          treatment  0.02927496 5.596401e-01  0.6112826
## 4 ben_past:treatment -0.03427496 5.305509e-01 -0.7077525
##
## $pgp_amount_end
##                 term  estimate     p.value  statistic
## 1        (Intercept) 308.59638 0.003833935  6.6735971
## 2           ben_past -19.69638 0.863392624 -0.2041236
## 3          treatment -45.10438 0.369729061 -0.9561169
## 4 ben_past:treatment  49.90438 0.654904093  0.4949424
```

```
benPast_pvalues <- sapply(benPastList, function(x) x$p.value)
benPast_tstat <- sapply(benPastList, function(x) x$statistic)
rownames(benPast_pvalues) <- c("intercept_p", "ben_past_p", "tr_p", "benpast:tr_p")
rownames(benPast_tstat) <- c("intercept_t", "ben_past_t", "tr_t", "benpast:tr_t")

mean(benPast_pvalues['benpast:tr_p',])
```

```
## [1] 0.5763561
```

```
mean(benPast_tstat['benpast:tr_t',])
```

```
## [1] -0.1788604
```

Comment 2

Fourth, I am concerned about the robustness of the findings given the relatively small sample size. In general, the findings were more mixed and more marginal than the paper's abstract and introduction would suggest. The marginality and mixed nature of the results is not surprising given the small sample size (20 communities), which is to be expected for such an intense/expensive intervention implemented in a real conflict setting. I think we can still learn something from the study, but would encourage the authors to address the following:

Were adjustments made for multiple comparisons? The PAP suggests that this would be done, but I do not believe that the manuscript mentions this.

Response Within family FDR (appendix). Footnote with appendix.

```
# from g-aggAnalysis-mainOutcomes.rmd
# comm
load("../survey_dat/d_analysis/list_of_coefs_and_ps.rda")
load("../obs_dat/b_analysis/obsDat_truePs.rda")

## add "base" var to PGG outcomes
newList[[6]]$base <- c(1,rep(0,3))
newList[[7]]$base <- c(1,rep(0,3))

## make list of ps
commListPs <- sapply(newList, function(x) x$truep[x$base %in% 1])
commListPs <- c(commListPs, obsDat_truePs$truep)
names(commListPs) <- c(outcome_list, rownames(obsDat_truePs))

# rebecca
## separate into our hypothesis families
### contact
p.adjust(commListPs[c(3,4,8:10)], method="none")
```

```
##             contactOnly                   rMean pastoralists_index_rank
##               0.0595000               0.2389000               0.0014000
##      farmers_index_rank     outgroup_index_rank
##               0.2416000               0.2534757
```

```
p.adjust(commListPs[c(3,4,8:10)], method="fdr")
```

```
##             contactOnly                   rMean pastoralists_index_rank
##               0.1487500               0.2534757               0.0070000
##      farmers_index_rank     outgroup_index_rank
##               0.2534757               0.2534757
```

```
p.adjust(commListPs[c(3,4,8:10)], method="holm")
```

```
##             contactOnly                   rMean pastoralists_index_rank
##                  0.2380                  0.7167                  0.0070
##      farmers_index_rank     outgroup_index_rank
##                  0.7167                  0.7167
```

```
### insecurity (its own family, could consider it an attitude)
p.adjust(commListPs[c(1,2,5)], method="none")
```

```
## attitude       in   end_exp
##   0.0448   0.0205    0.2117
```

```
p.adjust(commListPs[c(1,2,5)], method="fdr")
```

```
## attitude       in   end_exp
##   0.0672   0.0615    0.2117
```

```r
p.adjust(commListPs[c(1,2,5)], method="holm")
```

```
## attitude       in  end_exp
##   0.0896   0.0615   0.2117
```

```r
### attitudes
p.adjust(commListPs[c(1,5)], method="none")
```

```
## attitude  end_exp
##   0.0448   0.2117
```

```r
p.adjust(commListPs[c(1,5)], method="fdr")
```

```
## attitude  end_exp
##   0.0896   0.2117
```

```r
p.adjust(commListPs[c(1,5)], method="holm")
```

```
## attitude  end_exp
##   0.0896   0.2117
```

```r
### cooperation
p.adjust(commListPs[c(6,7)], method="none")
```

```
## pgp_donate_end pgp_amount_end
##         0.2938         0.8522
```

```r
p.adjust(commListPs[c(6,7)], method="fdr")
```

```
## pgp_donate_end pgp_amount_end
##         0.5876         0.8522
```

```r
p.adjust(commListPs[c(6,7)], method="holm")
```

```
## pgp_donate_end pgp_amount_end
##         0.5876         0.8522
```

```r
### all
p.adjust(p=commListPs, method = "none")
```

```
##                attitude                      in             contactOnly
##               0.0448000               0.0205000               0.0595000
##                   rMean                 end_exp          pgp_donate_end
##               0.2389000               0.2117000               0.2938000
##          pgp_amount_end pastoralists_index_rank      farmers_index_rank
##               0.8522000               0.0014000               0.2416000
##     outgroup_index_rank
##               0.2534757
```

```r
p.adjust(p=commListPs, method = "fdr")
```

```
##                attitude                       in              contactOnly
##               0.1487500                0.1025000                0.1487500
##                   rMean                  end_exp          pgp_donate_end
##               0.3168447                0.3168447                0.3264444
##          pgp_amount_end pastoralists_index_rank       farmers_index_rank
##               0.8522000                0.0140000                0.3168447
##     outgroup_index_rank
##               0.3168447
```

```r
p.adjust(p=commListPs, method = "holm")
```

```
##                attitude                       in              contactOnly
##                  0.3584                   0.1845                   0.4165
##                   rMean                  end_exp          pgp_donate_end
##                  1.0000                   1.0000                   1.0000
##          pgp_amount_end pastoralists_index_rank       farmers_index_rank
##                  1.0000                   0.0140                   1.0000
##     outgroup_index_rank
##                  1.0000
```

Individual-level (only interested in participant-control differences)

```r
## ind
load("../survey_dat/d_analysis/list_of_coefs_and_ps_ind.rda")

## add "base" var to PGG outcomes
newList_ind[[4]]$base <- c(rep(1,2),rep(0,6))
newList_ind[[5]]$base <- c(rep(1,2),rep(0,6))

# only interested in participant-control differences
indListPs <- sapply(newList_ind, function(x) x$truep[x$base %in% 1 & grepl("-part", rownames(x))])
names(indListPs) <- c(outcome_list[c(1:3,6:7)])

# not enough outcomes to really have families, but pgg vs not
# atts, insecurity, contact
p.adjust(p=indListPs[1:3], method = "none")
```

```
##    attitude          in contactOnly
##  0.13000000  0.18633333  0.01766667
```

```r
p.adjust(p=indListPs[1:3], method = "fdr")
```

```
##    attitude          in contactOnly
##   0.1863333   0.1863333   0.0530000
```

```r
p.adjust(p=indListPs[1:3], method = "holm")
```

```
##    attitude          in contactOnly
##       0.260       0.260       0.053
```

```r
# pgg
p.adjust(p=indListPs[4:5], method = "none")
```

```
## pgp_donate_end pgp_amount_end
##      0.2950000      0.8753333
```

```r
p.adjust(p=indListPs[4:5], method = "fdr")
```

```
## pgp_donate_end pgp_amount_end
##      0.5900000      0.8753333
```

```r
p.adjust(p=indListPs[4:5], method = "holm")
```

```
## pgp_donate_end pgp_amount_end
##      0.5900000      0.8753333
```

```r
### all
p.adjust(p=indListPs, method = "none")
```

```
##       attitude             in    contactOnly pgp_donate_end pgp_amount_end
##     0.13000000     0.18633333     0.01766667     0.29500000     0.87533333
```

```r
p.adjust(p=indListPs, method = "fdr")
```

```
##       attitude             in    contactOnly pgp_donate_end pgp_amount_end
##     0.31055556     0.31055556     0.08833333     0.36875000     0.87533333
```

```r
p.adjust(p=indListPs, method = "holm")
```

```
##       attitude             in    contactOnly pgp_donate_end pgp_amount_end
##     0.52000000     0.55900000     0.08833333     0.59000000     0.87533333
```

Comment 3

What are the substantive effect sizes and how do they compare to the minimum detectable effect sizes identified in the PAP's power analyses (0.3-0.4 SDs)?

Significant results are as strong or stronger than what is in the PAP Add power analysis to appendix Add .5 x-axis marker to community-level graph

Response *Power analysis showed we had ~0.80 power for*: (1) combined p-value of 0.3 SD for community and 0.20 SD for individual. (2) separate p-values of ~0.60 SD for community-level analysis (3) separate p-values of ~0.30-0.40 SD for individual-level analysis (this one done through egap power calc)

We would expect better power for our real analysis because the power analysis doesn't use randomization inference, but similar.

The effects we find are: *Comm-level* - close to 0.50 SD for survey outcomes attitudes, insecurity, contact - ~0.20 SD for survey experiments (endorsement, percent exp) - 1.0 SD for market pastoralists index - ~0.20 SDs for events outgroup index *Ind-level* - Between 0.15 and 0.35 SD for survey outcomes atts, insec, contact. Atts weakest, contact strongest.

```
# my original power analysis was bad. so inefficient and slow, so overcomplicated

# use our strongest outcome, insecurity.
#var=outcome_list_qip[2]; tau=0.2
bigPow.fn <-function(nsims, var=outcome_list_qip[2], tau)
{
  newPow.fn <- function(var, tau)
{
    # 6 TR sites from Nas, 4 from Ben
    newtr_nas <- sample(unique(ag.df$psu[ag.df$state %in% "nas"]), size=6)
    newtr_ben <- sample(unique(ag.df$psu[ag.df$state %in% "ben"]), size=4)
    newtr <- c(as.character(newtr_nas), as.character(newtr_ben))
    df <- ag.df
    df[,"newtr"] <- ifelse(df$psu %in% newtr, 1, 0)

  # make endline outcome with TR effect tau
  df[, paste0(var,"_end")] <- (df[,paste0(var,"_end")]-mean(df[,paste0(var,"_end")]))/sd(df[,paste0(var
  #scale(df[,paste0(var,"_end")])
  df[df$newtr %in% 1, paste0(var,"_end")] <- df[df$newtr %in% 1, paste0(var,"_end")]+tau

  # for baseline control, also scale
  df[, paste0(var,"_base")] <- (df[,paste0(var,"_base")]-mean(df[,paste0(var,"_base")]))/sd(df[,paste0(v

  # lm
  lm1 <- lm_robust(df[,paste0(var,"_end")]~df[,'newtr']+df[,paste0(var,"_base")]+state,
                   clusters = psu, data=df)

  want <- tidy(lm1)[2,5]
  return(want)
  }

  check <- do(nsims)*newPow.fn(var=var, tau=tau)
  pval <- mean(check<0.05)
  return(pval)
}
#bigPow.fn(nsims=100, tau=0.7)
```

Run it over taus from 0-1 for one survey outcome.

```
possibleTaus <- seq(0,1,0.1)
possibleTaus <- as.data.frame(possibleTaus)
system.time(
for(i in 1:nrow(possibleTaus))
{
  possibleTaus[i, "pow"] <- bigPow.fn(nsims=3000, tau=possibleTaus[i,1])
}
)
possibleTaus

#save
save(possibleTaus, file="new_power.Rdata")
```

```r
load("new_power.Rdata")
```

Comment 4

Figure 2 suggests that the effects in Figure 1 are driven by worsening conditions in control communities rather than improved relations in treatment communities. Given the small number of control communities (5), I worry that the effects could be driven by a particularly bad period in a single control community. Do similar patterns hold for other outcomes?

Response Great point. In general, we do see control communities getting worse across outcomes. But it wasn't driven by a particularly bad period in a single control community. Add to appendix: graphs with communities disaggregated

```r
# not the case that there is one big drop in control communities
## attitudes
summary(ag.df$attitude_cw[ag.df$treatment %in% 1])
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
## -0.52240 -0.10307  0.01343  0.01393  0.16472  0.40404
```

```r
summary(ag.df$attitude_cw[ag.df$treatment %in% 0])
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
## -0.33479 -0.21865 -0.12571 -0.08634  0.05927  0.14522
```

```r
## contact
summary(ag.df$contactOnly_cw[ag.df$treatment %in% 1])
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
## -0.62941 -0.11906  0.03868  0.03011  0.14413  0.57192
```

```r
summary(ag.df$contactOnly_cw[ag.df$treatment %in% 0])
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
## -0.66520 -0.24964 -0.11401 -0.10767  0.08487  0.45330
```

```r
## insecurity
summary(ag.df$in_cw[ag.df$treatment %in% 1])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -0.2871  0.1813  0.2407  0.2143  0.3228  0.4934
```

```r
summary(ag.df$in_cw[ag.df$treatment %in% 0])
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
## -0.17207 -0.08185 -0.04118  0.05516  0.22691  0.41297
```

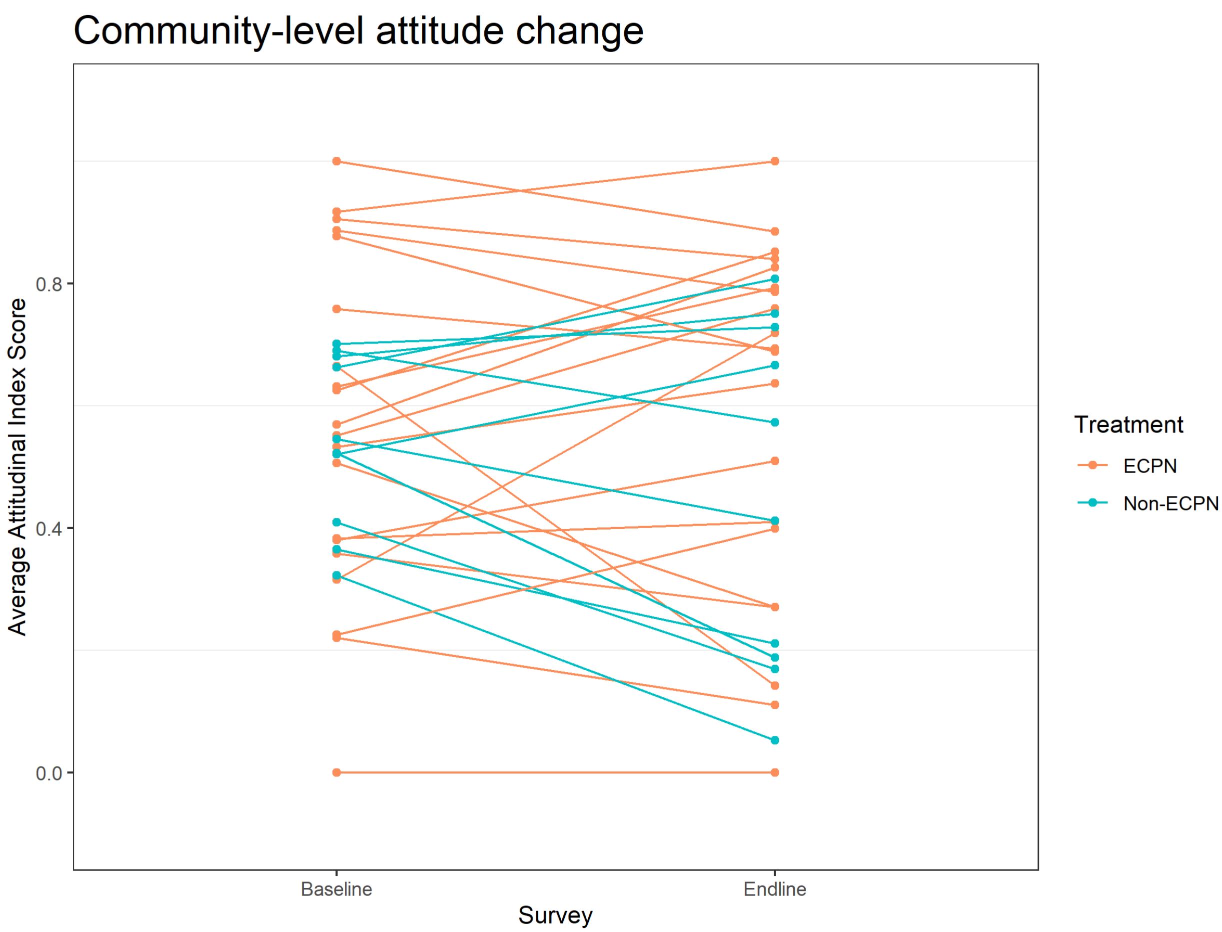

Figure 1: Each point represents a community. This graph shows (1) that overall changes are not driven by a large change in a single community and (2) that the overall change does not reflect a ceiling or floor effect.

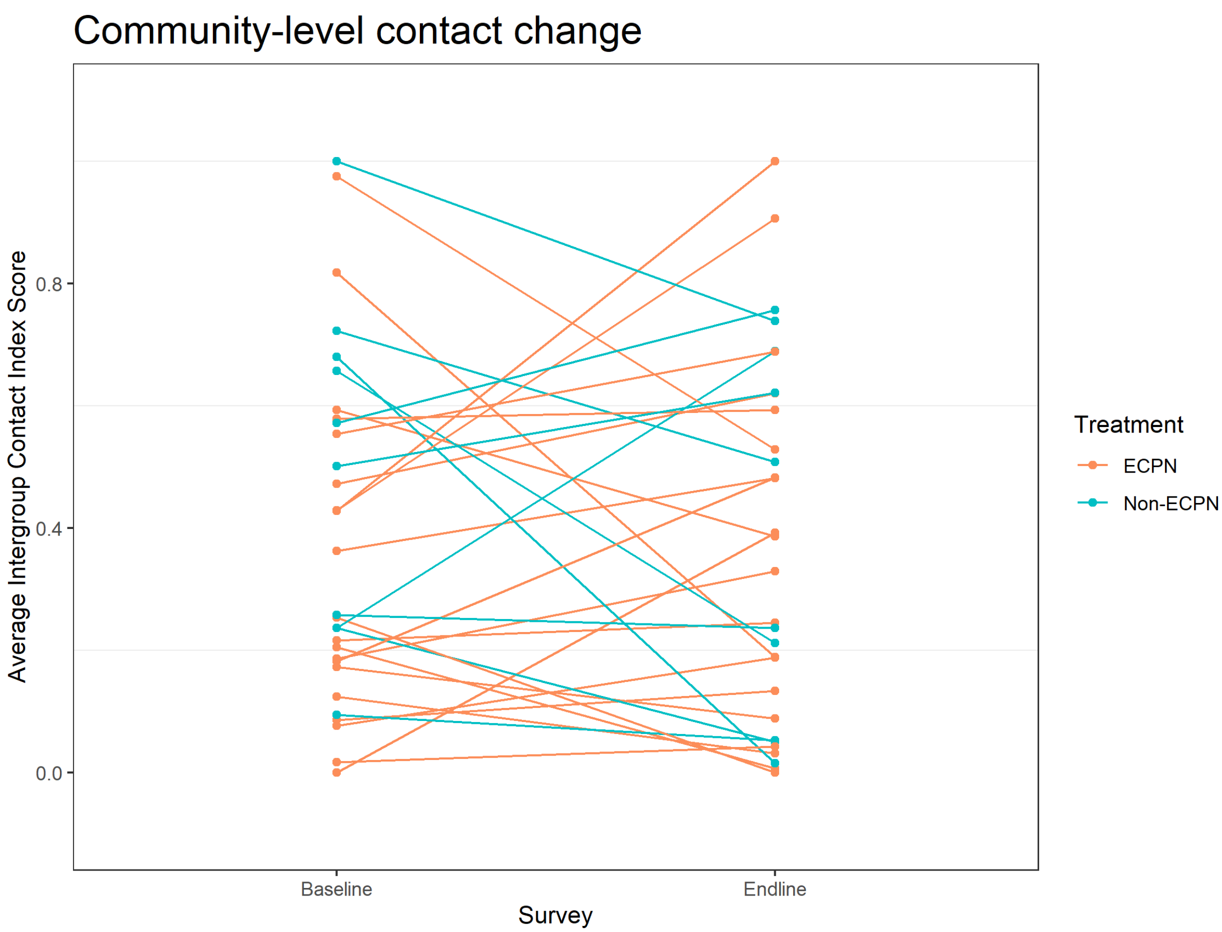

Figure 2: Each point represents a community. This graph shows (1) that overall changes are not driven by a large change in a single community and (2) that the overall change does not reflect a ceiling or floor effect.

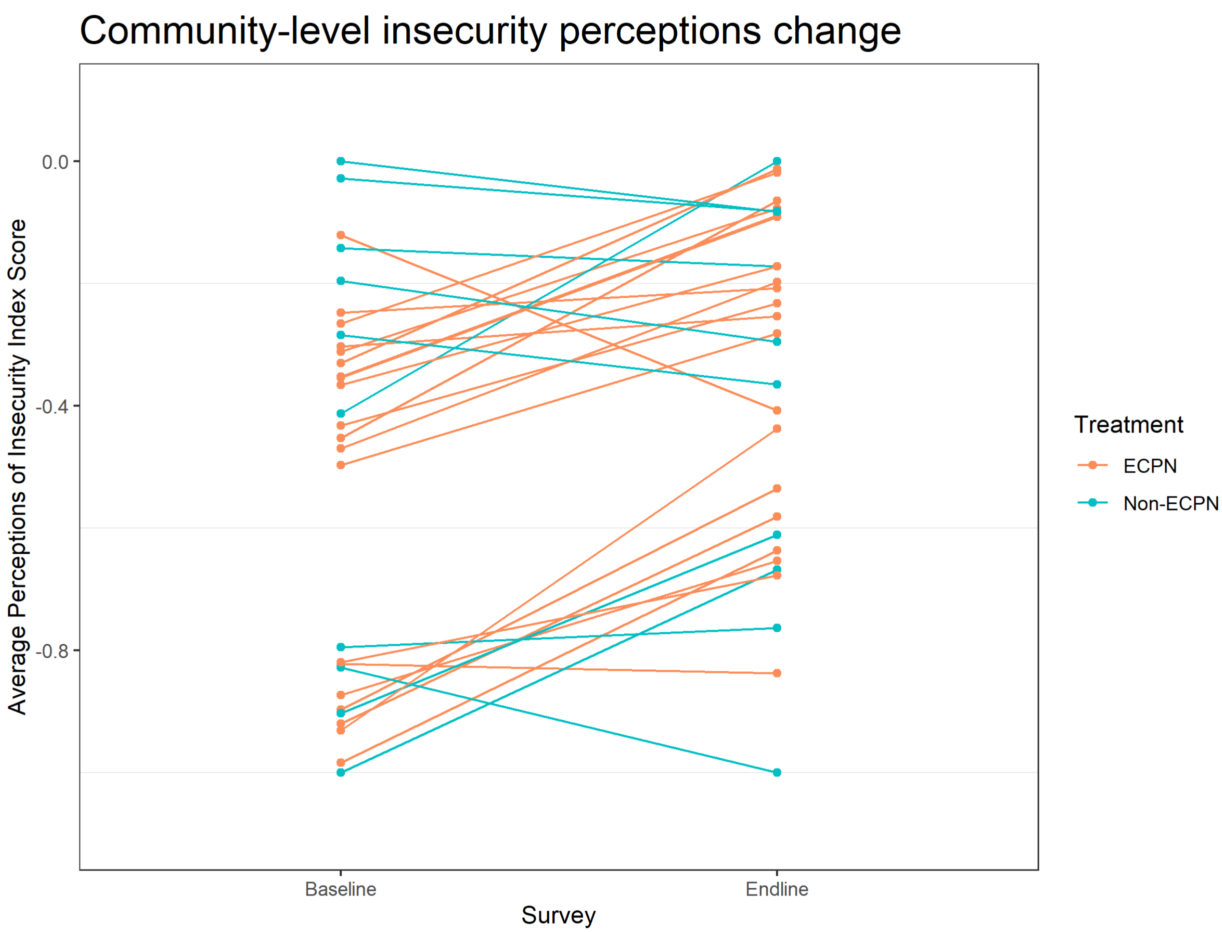

Figure 3: Each point represents a community. This graph shows (1) that overall changes are not driven by a large change in a single community and (2) that the overall change does not reflect a ceiling or floor effect.

Comment 5

Lots of the findings are not statistically significant (e.g., Figure 3 seems to show that none of the effects on non-participants were statistically distinguishable from zero). Sometimes such estimates are treated as null effects and other times as impacts that are not statistically significant. It would be helpful to understand how the authors make such distinctions.

Response Thank you, that is correct that the non-participants were not statistically different from the control group. We need to be consistent with wording.

.1 marginal .2 a potential trend />.2 null

```
# For exact p-values: Appendix B for community-level analysis, Appendix C for individual-level.
```

Finally, a few smaller questions: • Why are conflict and violence not included as dependent variables? There is an assumption that violence is caused by poor intergroup attitudes, but it is possible that attitudes could be improved without reducing violence. This seems important both theoretically and practically.

Response Good question. We have the footnote about it. It's a good idea but we weren't able to measure that. We have: violence_effect "In any clash that occurred in the last year, were you or anyone in your family negatively affected by an attack caused by [X group]"

- Were the enumerators who carried out the market observations aware of a community's treatment status? If such observations were not treatment-blind, should we be concerned about biased reporting of intergroup contact?

Response Lisa's email .

Method appendix Unaware of hypotheses Enumerator differences: take out Israel and Hadiza?

Review 2

Comment 1

The baseline for our understanding of pastoralist-farmer relations here refers to conflict. As the authors note circa line 147, relations had involved positive interactions that arose endogenously (see 167-68). The intervention here, as well as some of the drivers of contemporary conflict (ie, climate change) are largely exogenous. This could interfere with certain aspects of the research design, discussed below. But I believe it could be sufficiently, appropriately addressed within the framing of the paper around intergroup theory. For example, Allport was exploring prejudice and misinformed views about outgroups. Farmers and pastoralists have different interests, meaning that the attitudes shaping their interactions are not simply about a lack of information (or experience with) the other. The authors explore this around line 163. The PGG generates a shared interest, and in his regard the failure, or limited evidence of cooperation there, is an important limitation of this study, as the authors concede. I would appreciate a stronger justification, perhaps within the background about Nigeria itself, for how the inorganic and engineered contact does not interfere with the theorized drivers of outcomes. This seems especially important since participants were provided with mediation training; this made me wonder if I am observing effects and positive spillover of mediation training or random social contact that generates "authentic" attitudes and behaviors. Is the contact or the efficacy of the mediation driving change here?

Response Many good points being made here. We will address each of them.

Emphasize conditions need to be met. Why intervention was explicitly designed to meet those conditions. Just encountering can have negative effects (Enos) We can't disqualify that some of the spillover contact–for example in the markets–did not also contribute to outcomes. But that is part of the hope of contact interventions is that it leads to positive contact outside of the intervention. This is a rare study that we look at both in the intervention and outside (Mousa an exception).

Mediation being driving, see response above Lack of information–contact may apply in both when people have negative attitudes–extensions of contact theory to situations where knowledge may not be an issues but still negative beliefs

So in my rereading of the comment, it seems more about that in addition to lack of knowledge there are differing interests in this situation. So it's not just about knowledge–there is real resource scarcity. While the resource scarcity is real, a) it's a common interest to manage it well and b) there is a history of cooperation. Maybe Kertzer on contact not improving attitudes if other side wants bad things for you Maybe other stuff

Maybe concerned with: - artificiality of intervention? - real effect caused by random social contact that our intervention generated. Voluntary contact outside of the program is an intended outcome of the intervention. - these groups have differing interests, intervention trying to build on common interests. vs most contact interventions: cause of prejudice/conflict lack of knowledge. Here groups are competing with one another. Response Empirical question we are answering. beyond minimal groups, unrealistic that groups have no information. contact can do more than knowledge – can change perceptions about degree to which interests are aligned/misaligned. pettigrew & tropp meta-analysis about more than information.

Possible response solution: note that the reviewer makes good points and this is a complex issue. Then go over the several things we think reviewer is pointing out, responding to each point.

Comment 2

With regard to research design, I would appreciate a stronger justification of the site selection. As the authors note, Benue has embraced especially strong policies biased against pastoralists (it was one of the first states to ban "open grazing"). The intervention is therefore correcting for a manufactured imbalance between the groups, in a way: pastoralists likely feel insecure when the state sided with farmers and organized militias on their behalf. This is also a condition that is arguably different from Nasarawa. Perhaps this is adequately captured in the fixed effects models noted in the supplemental info. Secondly, the description of the group identities is standard, characterizing pastoralists as overwhelmingly Fulani and Muslim etc. There might be little reason to question this, but I noticed that individual characteristics are dropped from the survey analysis in the supplemental index… if I am understanding that reference properly.

Response In appendix, add exploratory regression that also uses individual's demographic characteristics Potentially add something about site selection or clarify the anti-grazing policies were implemented after site selection.

Misc

Power analysis

Chris: I tried to run a randomization inference power analysis. It takes too long to run.

But that will probably be conservative because it does not use randomization inference, so here it is with the rand inference function we used for p-values, but with tau added to endline outcomes. - simulate an effect: - make new treatment variable by shuffling site-level TR within states. - SD scale real baseline outcome - make new endline outcome by shuffling baseline outcome (within state) & adding tau - get true p-value - get sim_coef from lm(endline_outcome~newtr+baseline_outcome+state) - get null distribution - make null_tr by shuffling newtr at site-level, within states. - do nsims (e.g. 3000) regressions to make null distribution - compare sim_coef to null distribution: what % of coefs are bigger? - save true p-value: % of larger coefs - get power: how many true p-vals are below 0.05? - run true p-val function over several taus

```r
var=outcome_list_qip[2]; tr="treatment"
big_truePow.fn <-function(nsims, var=outcome_list_qip[2], tau){

truePow.fn <- function(var,tr,nsims=1000, dat=ag.df, tau=0)
{
  #(1) simulate an effect

  #make new treatment variable by shuffling site-level TR within states.
  ## real exp had 6 TR sites from Nas, 4 from Ben
  newtr_nas <- sample(unique(dat$psu[dat$state %in% "nas"]), size=6)
  newtr_ben <- sample(unique(dat$psu[dat$state %in% "ben"]), size=4)
  newtr <- c(as.character(newtr_nas), as.character(newtr_ben))
  dat[,"newtr"] <- ifelse(dat$psu %in% newtr, 1, 0)

  #SD scale real baseline outcome
  dat[, paste0(var,"_base")] <- (dat[,paste0(var,"_base")]-mean(dat[,paste0(var,"_base")]))/sd(dat[,pas

  #make new endline outcome by shuffling baseline outcome (within state) & adding tau
  dat <- dat %>% dplyr::group_by(state) %>%
    mutate(newout = sample(dat[, paste0(var,"_base")], replace=F)) %>%
    as.data.frame(.)
  dat$newout <- dat$newout + tau

  #(2) get true p-value

  #get sim_coef from lm(endline_outcome~newtr+baseline_outcome+state)
  thelm <- lm(newout~newtr+dat[, paste0(var,"_base")]+state, data=dat)
  thecoef <-coef(thelm)[2]

  #(2a) get null distribution
  null_dist = rep(NA,nsims)

  #make null_tr by shuffling newtr at site-level, within states.
  for(i in 1:nsims){
    nulltr_nas <- sample(unique(dat$psu[dat$state %in% "nas"]), size=6)
    nulltr_ben <- sample(unique(dat$psu[dat$state %in% "ben"]), size=4)
    nulltr <- c(as.character(nulltr_nas), as.character(nulltr_ben))
    dat[,"nulltr"] <- ifelse(dat$psu %in% nulltr, 1, 0)

    #do nsims (e.g. 3000) regressions to make null distribution
    null_lm <- lm(newout~nulltr+dat[, paste0(var,"_base")]+state, data=dat)
    null_dist[i] <- coef(null_lm)[2]
  }

  #compare sim_coef to null distribution: what % of coefs are bigger?
  thep <- mean(null_dist<thecoef)
  #true p-value == % of larger coefs
  return(thep)
}

  # get power: how many true p-vals are below 0.05?
  check <- do(nsims)*truePow.fn(var=var, tau=tau)
  pval <- mean(check<0.05)
```

```
  return(pval)

}
```

Run over taus. chris: nevermind, this didn't run in 3 hours. Based on how long it takes to run 10 times, if time scales linearly it will run for 166 hours. Non randomization inference power analysis is fine.

```
possibleTaus <- seq(0,1,0.1)
possibleTaus <- as.data.frame(possibleTaus)
system.time(
for(i in 1:nrow(possibleTaus))
{
  possibleTaus[i, "pow"] <- big_truePow.fn(nsims=1000, tau=possibleTaus[i,1])
}
)
possibleTaus

#save
save(possibleTaus, file="new_power_truep.Rdata")
```

```
load("new_power_truep.Rdata")
```