

# Problem Set 3

Zachary Brandt

2025-11-15

## Weather and Witch Killing

Construct a new variable for the total number of murders in a village-year (witch + non-witch murders).

```
killing$total_murders <- killing$witch_murders + killing$oth_murders
```

Create a table of summary statistics for all variables in the dataset, including the mean, standard deviation, minimum, maximum, and number of observations, using stargazer, summary or describe commands in R. Discuss any noteworthy patterns. Pay particular attention to the murder and rainfall variables.

```
summary(killing)
```

```
##      vid          year    witch_murders     oth_murders
##  Min.   : 1.00   Min.   :1992   Min.   :0.00000   Min.   :0.00000
##  1st Qu.:17.00  1st Qu.:1994  1st Qu.:0.00000  1st Qu.:0.00000
##  Median :34.00  Median :1997  Median :0.00000  Median :0.00000
##  Mean   :35.03  Mean   :1997  Mean   :0.09103  Mean   :0.09103
##  3rd Qu.:54.00  3rd Qu.:2000 3rd Qu.:0.00000  3rd Qu.:0.00000
##  Max.   :71.00  Max.   :2002  Max.   :3.00000  Max.   :5.00000
##      any_rain    any_disease    famine       educat
##  Min.   :0.0000  Min.   :0.0000  Min.   :0.0000  Min.   :0.8571
##  1st Qu.:0.0000  1st Qu.:0.0000  1st Qu.:0.0000  1st Qu.:3.5263
##  Median :0.0000  Median :0.0000  Median :0.0000  Median :4.2000
##  Mean   :0.1712  Mean   :0.1481  Mean   :0.1739  Mean   :4.0346
##  3rd Qu.:0.0000  3rd Qu.:0.0000  3rd Qu.:0.0000  3rd Qu.:4.6875
##  Max.   :1.0000  Max.   :1.0000  Max.   :1.0000  Max.   :6.6667
##      trad_relig    total_murders
##  Min.   :0.0000  Min.   :0.0000
##  1st Qu.:0.5333  1st Qu.:0.0000
##  Median :0.6500  Median :0.0000
##  Mean   :0.6541  Mean   :0.1821
##  3rd Qu.:0.8000  3rd Qu.:0.0000
##  Max.   :1.0000  Max.   :5.0000
```

Now consider the effect of extreme weather on murders in the village.

Install “miceadds” and “sandwich”. Using the lm.cluster command, regress total murders (in a village in a particular year) on the indicator for whether a drought or flood occurred in that year. Make sure that error terms should be allowed to be correlated (“clustered”) across years for the same village (use vid). Simply use summary to report the results in this question. [Note: Results estimated by lm.cluster could not be exported directly with stargazer so we use summary for simplicity. In the section we will teach how to export clustered regression results in a neater way.]

```
library(miceadds)
library(sandwich)
```

```
model <- lm.cluster(data = killing, total_murders ~ any_rain, cluster = "vid")
summary(model)
```

```
## R^2= 0.00125
##
##             Estimate Std. Error    t value   Pr(>|t|)
## (Intercept) 0.17377049 0.02208164 7.869457 3.561844e-15
## any_rain     0.04845173 0.04567845 1.060713 2.888203e-01
```

In a second regression, add average years of schooling and proportion of households practicing traditional religions as additional explanatory variables.

```
model <- lm.cluster(data = killing, total_murders ~ any_rain + educat +
                     trad_relig, cluster = "vid")
summary(model)
```

```
## R^2= 0.00738
##
##             Estimate Std. Error    t value   Pr(>|t|)
## (Intercept) 0.3276266665 0.14575336 2.247815515 0.02458796
## any_rain     0.0400209562 0.04263900 0.938599790 0.34793626
## educat      -0.0378694872 0.02644472 -1.432024384 0.15213684
## trad_relig   0.0005751421 0.10422490 0.005518279 0.99559707
```

Finally, consider a possible instrumental variables (IV) approach. Economic theory suggests that extreme economic hardship-such as a famine-may be associated with more violence, including murders. Famine may be caused by extreme rainfall (which would be the instrumental variable).

Write out the first stage regression, the second stage regression, and the reduced form regression.

```
# First stage regression
any_rain ~ famine

# Second stage regression
total_murders ~ any_rain + educat + trad_relig

# Reduced form regression
total_murders ~ famine + educat + trad_relig
```

## The Primary School Deworming Project (PSDP)

Use R to estimate regressions in the format of the following. Simply use summary to report the results in this question. To receive full credits, please highlight the names of dependent variables and estimated coefficients of treatment with red rectangles. You could do this by annotating the pdf document compiled from .rmd or exported from Microsoft Word.

name\_dep\_var = treatment + name\_control\_vars, with “weight” as the sampling weight and “psdpsch98” as the cluster ID

```
model <- lm.cluster(data = deworming,
                      formula = as.formula(paste("totyrs enrolled" ~ treatment +", controls)),
                      cluster = "psdpsch98",
                      weights = deworming$weight)
summary(model)

## R^2= 0.29307
##
##             Estimate Std. Error    t value   Pr(>|t|)
```

```

## (Intercept)      9.36193350  0.23062645  40.59349415  0.000000e+00
## treatment       0.29293848  0.14457838  2.02615697  4.274870e-02
## saturation_dm   1.04973111  0.83687238  1.25435029  2.097147e-01
## demeaned_popT_6k 0.11471369  0.04544937  2.52398868  1.160317e-02
## zoneidI2        -0.35050682  0.29689217 -1.18058626  2.377671e-01
## zoneidI3        -0.33705209  0.26835345 -1.25600059  2.091158e-01
## zoneidI4        -0.71992160  0.24136882 -2.98266197  2.857534e-03
## zoneidI5        -0.34618274  0.18964652 -1.82541046  6.793911e-02
## zoneidI6        -0.33159907  0.25912715 -1.27967703  2.006587e-01
## zoneidI7        -0.18222813  0.22504363 -0.80974577  4.180863e-01
## zoneidI8        -0.29463521  0.20355627 -1.44743859  1.477741e-01
## pup_pop          0.05735229  0.31829823  0.18018414  8.570080e-01
## month_interviewI2 -0.05851750  0.20532211 -0.28500338  7.756416e-01
## month_interviewI3 -0.26699350  0.24919370 -1.07142954  2.839763e-01
## month_interviewI4  0.06983719  0.21985684  0.31764847  7.507516e-01
## month_interviewI5  0.06284821  0.18252743  0.34432201  7.306041e-01
## month_interviewI6 -0.52084524  0.21061417 -2.47298284  1.339906e-02
## month_interviewI7 -0.12045817  0.21799893 -0.55256314  5.805626e-01
## month_interviewI8 -0.09906835  0.19625913 -0.50478341  6.137110e-01
## month_interviewI9 -0.55896077  0.19810614 -2.82152174  4.779640e-03
## month_interviewI10 -0.41672577  0.21515277 -1.93688309  5.275964e-02
## month_interviewI11  0.01662161  0.21979527  0.07562316  9.397189e-01
## month_interviewI12  0.20870572  0.21561415  0.96795929  3.330647e-01
## cost_sharing     -0.25515415  0.13983446 -1.82468728  6.804823e-02
## std98_base_I2    -0.83988629  0.13756892 -6.10520385  1.026696e-09
## std98_base_I3    -1.63654576  0.13385499 -12.22625893  2.250676e-34
## std98_base_I4    -2.43389048  0.14885867 -16.35034418  4.325301e-60
## std98_base_I5    -3.53685841  0.14889959 -23.75331158  1.015487e-124
## std98_base_I6    -4.22771777  0.17284421 -24.45970073  3.968030e-132
## female_baseline  -0.84317914  0.10269801 -8.21027752  2.206778e-16
## avgtest96        0.62820130  0.22734032  2.76326386  5.722648e-03

model <- lm.cluster(data = deworming,
                      formula = as.formula(paste("passed_primary_exam ~ treatment +", controls)),
                      cluster = "psdpsch98",
                      weights = deworming$weight)
summary(model)

## R^2= 0.07025
##
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)            0.55703765 0.041667729 13.3685628 9.230720e-41
## treatment              0.05059504 0.030832345  1.6409727 1.008031e-01
## saturation_dm          0.21401952 0.161796316  1.3227713 1.859115e-01
## demeaned_popT_6k        0.00648825 0.009623449  0.6742126 5.001762e-01
## zoneidI2                -0.01851791 0.053875838 -0.3437147 7.310609e-01
## zoneidI3                -0.02933843 0.056266748 -0.5214168 6.020765e-01
## zoneidI4                -0.06412930 0.041452900 -1.5470401 1.218536e-01
## zoneidI5                -0.04823160 0.041268200 -1.1687351 2.425104e-01
## zoneidI6                -0.09347771 0.044252900 -2.1123522 3.465626e-02
## zoneidI7                -0.01260101 0.039584380 -0.3183329 7.502325e-01
## zoneidI8                -0.03444625 0.035483581 -0.9707660 3.316648e-01
## pup_pop                 -0.04467231 0.057779367 -0.7731534 4.394316e-01
## month_interviewI2       -0.02470522 0.037212977 -0.6638873 5.067624e-01
## month_interviewI3       0.01908657 0.046188352  0.4132333 6.794357e-01

```

```

## month_interviewI4    0.02680356 0.045131554  0.5938984 5.525801e-01
## month_interviewI5    0.07783873 0.034265023  2.2716673 2.310662e-02
## month_interviewI6   -0.03233291 0.035497349 -0.9108542 3.623722e-01
## month_interviewI7    0.05702830 0.035993701  1.5843966 1.131035e-01
## month_interviewI8    0.01787235 0.034023825  0.5252894 5.993820e-01
## month_interviewI9   -0.06583911 0.033295183 -1.9774366 4.799230e-02
## month_interviewI10  -0.01787251 0.037135018 -0.4812845 6.303143e-01
## month_interviewI11    0.03459020 0.037483934  0.9228007 3.561111e-01
## month_interviewI12    0.08543936 0.044767967  1.9084932 5.632750e-02
## cost_sharing        -0.03940312 0.026995349 -1.4596263 1.443928e-01
## std98_base_I2        0.04144177 0.030373615  1.3644004 1.724416e-01
## std98_base_I3        0.09995566 0.034478228  2.8990951 3.742413e-03
## std98_base_I4        0.09042610 0.027376615  3.3030418 9.564214e-04
## std98_base_I5        0.09483187 0.031264189  3.0332426 2.419410e-03
## std98_base_I6        0.08928989 0.026568670  3.3607210 7.773930e-04
## female_baseline     -0.18346549 0.019210568 -9.5502375 1.293993e-21
## avgtest96           0.13895945 0.039079567  3.5558083 3.768187e-04

model <- lm.cluster(data = deworming,
                      formula = as.formula(paste("num_meals_yesterday" ~ treatment +, controls)),
                      cluster = "psdpsch98",
                      weights = deworming$weight)
summary(model)

## R^2= 0.03368
##
##                               Estimate Std. Error      t value  Pr(>|t|)
## (Intercept)                2.137800535 0.054071596 39.5364793 0.000000e+00
## treatment                  0.095214738 0.028534499  3.3368288 8.474014e-04
## saturation_dm              0.404266462 0.125253242  3.2275928 1.248365e-03
## demeaned_popT_6k            -0.014600039 0.009440756 -1.5464905 1.219862e-01
## zoneidI2                  -0.206007638 0.041904835 -4.9160828 8.829306e-07
## zoneidI3                  -0.105488852 0.048372829 -2.1807460 2.920221e-02
## zoneidI4                  -0.182413806 0.049443528 -3.6893363 2.248398e-04
## zoneidI5                  -0.040890385 0.035695649 -1.1455286 2.519903e-01
## zoneidI6                  -0.160519728 0.054835676 -2.9272864 3.419339e-03
## zoneidI7                  -0.155582782 0.054927217 -2.8325262 4.618178e-03
## zoneidI8                  -0.117487515 0.057083089 -2.0581843 3.957245e-02
## pup_pop                    -0.129344800 0.056389843 -2.2937606 2.180425e-02
## month_interviewI2          -0.005970265 0.058303488 -0.1023998 9.184393e-01
## month_interviewI3          -0.070618194 0.045376447 -1.5562742 1.196429e-01
## month_interviewI4          -0.019954171 0.054799429 -0.3641310 7.157602e-01
## month_interviewI5          0.084093784 0.057604882  1.4598378 1.443347e-01
## month_interviewI6          0.108489699 0.053660956  2.0217623 4.320091e-02
## month_interviewI7          0.109581708 0.061996822  1.7675375 7.713825e-02
## month_interviewI8          0.080570005 0.046600183  1.7289633 8.381567e-02
## month_interviewI9          0.068880307 0.068668539  1.0030839 3.158204e-01
## month_interviewI10         0.068954527 0.053630068  1.2857438 1.985325e-01
## month_interviewI11         0.093683454 0.052829555  1.7733152 7.617649e-02
## month_interviewI12         0.094273287 0.056944788  1.6555209 9.781888e-02
## cost_sharing               -0.072602514 0.031463293 -2.3075307 2.102525e-02
## std98_base_I2              0.095995229 0.036459465  2.6329303 8.465172e-03
## std98_base_I3              0.099110690 0.025664884  3.8617237 1.125899e-04
## std98_base_I4              0.115604211 0.041040041  2.8168639 4.849506e-03
## std98_base_I5              0.094638757 0.037128737  2.5489355 1.080523e-02

```

```

## std98_base_I6      0.077599439 0.034085379  2.2766195 2.280896e-02
## female_baseline   0.078479802 0.025786968  3.0433900 2.339289e-03
## avgtest96         0.047527477 0.030363191  1.5652991 1.175128e-01

# Regression 4: Total hours worked
model <- lm.cluster(data = deworming,
                      formula = as.formula(paste("total_hours ~ treatment +", controls)),
                      cluster = "psdpsch98",
                      weights = deworming$weight)
summary(model)

## R^2= 0.05921
##
##                               Estimate Std. Error     t value    Pr(>|t|)
## (Intercept)                13.08876272  2.3381078  5.59801508 2.168199e-08
## treatment                  1.59859758  1.0363162  1.54257702 1.229334e-01
## saturation_dm              9.67442053  7.8239659  1.23651109 2.162686e-01
## demeaned_popT_6k            0.17674733  0.3611574  0.48939143 6.245646e-01
## zoneidI2                   1.17502130  2.3036805  0.51006261 6.100076e-01
## zoneidI3                   0.54679120  2.3153729  0.23615687 8.133109e-01
## zoneidI4                   2.80939706  2.2247462  1.26279441 2.066631e-01
## zoneidI5                   1.47963117  1.9343260  0.76493371 4.443110e-01
## zoneidI6                   3.19678682  1.8918127  1.68980087 9.106606e-02
## zoneidI7                   -0.47977600  2.0075523 -0.23898555 8.111168e-01
## zoneidI8                   1.15439311  1.9203894  0.60112450 5.477571e-01
## pup_pop                     -2.15632733  2.8595008 -0.75409224 4.507938e-01
## month_interviewI2           0.59573594  2.3590944  0.25252739 8.006334e-01
## month_interviewI3           -0.06585665  1.6962492 -0.03882487 9.690300e-01
## month_interviewI4           2.41820950  2.1724916  1.11310420 2.656637e-01
## month_interviewI5           0.46863387  1.8747642  0.24996950 8.026109e-01
## month_interviewI6           -0.14118560  1.7967851 -0.07857679 9.373693e-01
## month_interviewI7           0.26229886  1.8579415  0.14117714 8.877300e-01
## month_interviewI8           2.70864575  1.6499411  1.64166212 1.006600e-01
## month_interviewI9           3.39383629  1.7666636  1.92104270 5.472632e-02
## month_interviewI10          4.38247383  1.9968526  2.19469068 2.818579e-02
## month_interviewI11          2.29345736  1.6612186  1.38058732 1.674059e-01
## month_interviewI12          2.61232134  2.3204101  1.12580155 2.602495e-01
## cost_sharing                -1.58852908  0.8426322 -1.88519863 5.940303e-02
## std98_base_I2                2.87021252  1.1831510  2.42590540 1.527025e-02
## std98_base_I3                5.25419541  1.4098919  3.72666536 1.940298e-04
## std98_base_I4                9.11589669  1.6967973  5.37241346 7.768972e-08
## std98_base_I5                11.75981169  1.3679153  8.59688610 8.190821e-18
## std98_base_I6                10.43739045  1.4692341  7.10396673 1.212261e-12
## female_baseline               -6.63145907  0.9455658 -7.01321801 2.328981e-12
## avgtest96                    -0.44322869  1.5109010 -0.29335389 7.692517e-01

model <- lm.cluster(data = deworming,
                      formula = as.formula(paste("ln_emp_salary_total ~ treatment +", controls)),
                      cluster = "psdpsch98",
                      weights = deworming$weight)
summary(model)

## R^2= 0.18863
##
##                               Estimate Std. Error     t value    Pr(>|t|)

```

```

## (Intercept) 7.639928289 0.19773326 38.63754808 0.000000e+00
## treatment 0.264505216 0.08519794 3.10459639 1.905388e-03
## saturation_dm 0.946135127 0.86137458 1.09840149 2.720292e-01
## demeaned_popT_6k 0.032530501 0.02671009 1.21791074 2.232579e-01
## zoneidI2 -0.182680139 0.14830549 -1.23178272 2.180303e-01
## zoneidI3 -0.167813300 0.15061279 -1.11420348 2.651919e-01
## zoneidI4 -0.142579282 0.14632985 -0.97436906 3.298733e-01
## zoneidI5 -0.111112127 0.18852035 -0.58939062 5.555993e-01
## zoneidI6 0.129125628 0.13982088 0.92350747 3.557428e-01
## zoneidI7 -0.119284338 0.14656991 -0.81383918 4.157371e-01
## zoneidI8 -0.070609931 0.18902352 -0.37355104 7.087384e-01
## pup_pop -0.004367281 0.25788153 -0.01693522 9.864883e-01
## month_interviewI2 -0.424540364 0.20029479 -2.11957764 3.404168e-02
## month_interviewI3 -0.248233855 0.19172212 -1.29475857 1.954035e-01
## month_interviewI4 -0.340358635 0.25542149 -1.33253721 1.826837e-01
## month_interviewI5 -0.146851797 0.16768340 -0.87576823 3.811560e-01
## month_interviewI6 -0.115661602 0.15892609 -0.72776975 4.667545e-01
## month_interviewI7 0.137806747 0.14837657 0.92876355 3.530116e-01
## month_interviewI8 -0.009137235 0.14447234 -0.06324557 9.495710e-01
## month_interviewI9 -0.063554020 0.13255454 -0.47945561 6.316145e-01
## month_interviewI10 -0.080449366 0.12892352 -0.62400843 5.326220e-01
## month_interviewI11 -0.218388053 0.15388675 -1.41914782 1.558559e-01
## month_interviewI12 0.145665489 0.18990377 0.76704895 4.430524e-01
## cost_sharing -0.148538108 0.08788652 -1.69011257 9.100642e-02
## std98_base_I2 0.042169073 0.15448271 0.27296952 7.848766e-01
## std98_base_I3 0.327266912 0.15126070 2.16359517 3.049543e-02
## std98_base_I4 0.549476543 0.17078767 3.21730799 1.293996e-03
## std98_base_I5 0.576142925 0.14800390 3.89275491 9.911227e-05
## std98_base_I6 0.706708226 0.17421753 4.05647025 4.981990e-05
## female_baseline -0.436652085 0.08992569 -4.85569896 1.199628e-06
## avgtest96 0.032740253 0.12304016 0.26609404 7.901668e-01

```

Deworming benefits might be stronger for certain groups — for instance, girls (perhaps because they were more likely to be infected) or children with lower BMI at baseline (because they were less healthy initially).

Please estimate whether the deworming treatment had a differential impact on totyrs\_enrolled, passed\_primary\_exam, and total\_hours by gender (female\_baseline) and then by BMI (BMI). To receive full credits, please highlight the names of dependent variables and estimated coefficients of interactive terms with red rectangles.

```

model <- lm.cluster(data = deworming,
                      formula = as.formula(paste("totyrs_enrolled" ~ treatment * female_baseline +",
                      cluster = "psdpsch98",
                      weights = deworming$weight))
summary(model)

## R^2= 0.29309
##
##                               Estimate Std. Error      t value      Pr(>|t|)
## (Intercept) 9.34001706 0.24721368 37.78115022 2.317739e-312
## treatment 0.32385023 0.18053697 1.79381668 7.284243e-02
## female_baseline -0.79960859 0.17844147 -4.48106943 7.426994e-06
## saturation_dm 1.05224345 0.83649024 1.25792675 2.084182e-01
## demeaned_popT_6k 0.11466494 0.04546365 2.52212371 1.166487e-02
## zoneidI2 -0.34974800 0.29800975 -1.17361260 2.405502e-01
## zoneidI3 -0.33726921 0.26900407 -1.25376990 2.099257e-01

```

```

## zoneidI4          -0.72133229 0.24104169 -2.99256231 2.766462e-03
## zoneidI5          -0.34623462 0.18997892 -1.82248965 6.838072e-02
## zoneidI6          -0.33210664 0.25915253 -1.28151032 2.000145e-01
## zoneidI7          -0.18283586 0.22464763 -0.81387840 4.157146e-01
## zoneidI8          -0.29404148 0.20403279 -1.44114815 1.495428e-01
## pup_pop            0.05876268 0.31944114 0.18395463 8.540490e-01
## month_interviewI2 -0.05861037 0.20542269 -0.28531598 7.754021e-01
## month_interviewI3 -0.26636026 0.24879319 -1.07060911 2.843452e-01
## month_interviewI4  0.06941129 0.22006370 0.31541454 7.524469e-01
## month_interviewI5  0.06263877 0.18278430 0.34269226 7.318300e-01
## month_interviewI6  -0.52011182 0.21024141 -2.47387905 1.336549e-02
## month_interviewI7  -0.12064488 0.21824338 -0.55279972 5.804006e-01
## month_interviewI8  -0.09862315 0.19608533 -0.50296036 6.149921e-01
## month_interviewI9  -0.55875945 0.19805360 -2.82125365 4.783636e-03
## month_interviewI10 -0.41503373 0.21531819 -1.92753684 5.391276e-02
## month_interviewI11 0.01586017 0.21953159 0.07224552 9.424065e-01
## month_interviewI12 0.20674065 0.21657115 0.95460845 3.397757e-01
## cost_sharing       -0.25594858 0.13972764 -1.83176779 6.698602e-02
## std98_base_I2      -0.83902395 0.13697869 -6.12521504 9.056121e-10
## std98_base_I3      -1.63610041 0.13445383 -12.16849222 4.574545e-34
## std98_base_I4      -2.43318459 0.14785632 -16.45641291 7.544072e-61
## std98_base_I5      -3.53668628 0.14865619 -23.79104696 4.134304e-125
## std98_base_I6      -4.22676043 0.17291273 -24.44447293 5.761752e-132
## avgtest96          0.62918271 0.22797704 2.75985119 5.782770e-03
## treatment:female_baseline -0.06432997 0.21798401 -0.29511328 7.679073e-01

model <- lm.cluster(data = deworming,
                      formula = as.formula(paste("passed_primary_exam ~ treatment * female_baseline +",
                      cluster = "psdpsch98",
                      weights = deworming$weight))

summary(model)

## R^2= 0.07025
##
##                               Estimate Std. Error   t value Pr(>|t|)
## (Intercept)                0.557441964 0.043731353 12.74696345 3.241263e-37
## treatment                  0.050026456 0.033311587  1.50177346 1.331556e-01
## female_baseline             -0.184271956 0.031250938 -5.89652569 3.712350e-09
## saturation_dm               0.213988044 0.161758584  1.32288524 1.858735e-01
## demeaned_popT_6k              0.006490048 0.009610316  0.67532101 4.994718e-01
## zoneidI2                   -0.018527494 0.053867896 -0.34394316 7.308890e-01
## zoneidI3                   -0.029327980 0.056217244 -0.52169011 6.018861e-01
## zoneidI4                   -0.064099432 0.041399314 -1.54832113 1.215450e-01
## zoneidI5                   -0.048223975 0.041252990 -1.16898134 2.424112e-01
## zoneidI6                   -0.093467000 0.044194572 -2.11489770 3.443867e-02
## zoneidI7                   -0.012589094 0.039497033 -0.31873518 7.499273e-01
## zoneidI8                   -0.034456278 0.035536450 -0.96960381 3.322440e-01
## pup_pop                     -0.044706478 0.057794243 -0.77354553 4.391996e-01
## month_interviewI2            -0.024703224 0.037210926 -0.66387016 5.067734e-01
## month_interviewI3            0.019074646 0.046241015  0.41250492 6.799694e-01
## month_interviewI4            0.026810191 0.045158556  0.59369018 5.527194e-01
## month_interviewI5            0.077843332 0.034248830  2.27287567 2.303367e-02
## month_interviewI6            -0.032345065 0.035473801 -0.91180148 3.618732e-01
## month_interviewI7            0.057033818 0.036001736  1.58419631 1.131491e-01
## month_interviewI8            0.017865714 0.034033036  0.52495212 5.996165e-01

```

```

## month_interviewI9      -0.065845139 0.033276147 -1.97874889 4.784429e-02
## month_interviewI10     -0.017905661 0.037223800 -0.48102721 6.304972e-01
## month_interviewI11      0.034604214 0.037522358  0.92222919 3.564090e-01
## month_interviewI12      0.085470424 0.044839977  1.90612103 5.663450e-02
## cost_sharing          -0.039386937 0.026927025 -1.46272889 1.435416e-01
## std98_base_I2           0.041424883 0.030331197  1.36575167 1.720169e-01
## std98_base_I3           0.099946134 0.034529616  2.89450464 3.797573e-03
## std98_base_I4           0.090411934 0.027279029  3.31433838 9.186027e-04
## std98_base_I5           0.094828133 0.031262793  3.03325855 2.419282e-03
## std98_base_I6           0.089267187 0.026633066  3.35174279 8.030459e-04
## avgtest96                0.138941993 0.039026271  3.56021700 3.705485e-04
## treatment:female_baseline 0.001190195 0.039931916  0.02980561 9.762221e-01

model <- lm.cluster(data = deworming,
                      formula = as.formula(paste("total_hours" ~ treatment * female_baseline +",
                      cluster = "psdpsch98",
                      weights = deworming$weight))

summary(model)

## R^2= 0.06075

##                                     Estimate Std. Error   t value Pr(>|t|)
## (Intercept)                  11.73503812 2.3874425  4.91531764 8.863863e-07
## treatment                     3.50617266 1.4796274  2.36963222 1.780579e-02
## female_baseline                -3.93672015 1.7186783 -2.29055086 2.198940e-02
## saturation_dm                  9.80617283 7.7735236  1.26148365 2.071346e-01
## demeaned_popT_6k                 0.17268152 0.3560491  0.48499349 6.276810e-01
## zoneidI2                      1.22186463 2.2477500  0.54359454 5.867205e-01
## zoneidI3                      0.53660107 2.2824402  0.23509973 8.141313e-01
## zoneidI4                      2.72192965 2.2250853  1.22329226 2.212193e-01
## zoneidI5                      1.47190779 1.9176250  0.76756809 4.427438e-01
## zoneidI6                      3.16812913 1.8665723  1.69729784 8.964037e-02
## zoneidI7                      -0.51508605 1.9869419 -0.25923559 7.954535e-01
## zoneidI8                      1.19552270 1.8866102  0.63368825 5.262843e-01
## pup_pop                       -2.07066353 2.8274408 -0.73234549 4.639577e-01
## month_interviewI2                  0.59056737 2.3502315  0.25128051 8.015972e-01
## month_interviewI3                 -0.02577300 1.6830379 -0.01531338 9.877822e-01
## month_interviewI4                  2.39506422 2.1786507  1.09933371 2.716225e-01
## month_interviewI5                  0.45658028 1.8761288  0.24336297 8.077242e-01
## month_interviewI6                 -0.09547212 1.7888885 -0.05336952 9.574375e-01
## month_interviewI7                  0.25580693 1.8506416  0.13822608 8.900617e-01
## month_interviewI8                  2.73216078 1.6349465  1.67110102 9.470173e-02
## month_interviewI9                  3.41292449 1.7546443  1.94508058 5.176528e-02
## month_interviewI10                 4.48770917 2.0103318  2.23232269 2.559364e-02
## month_interviewI11                 2.24371616 1.6676717  1.34541836 1.784902e-01
## month_interviewI12                 2.49346436 2.3349363  1.06789395 2.855684e-01
## cost_sharing                   -1.63622563 0.8293350 -1.97293681 4.850276e-02
## std98_base_I2                   2.92273160 1.1932968  2.44929147 1.431376e-02
## std98_base_I3                   5.28265334 1.4283013  3.69855677 2.168289e-04
## std98_base_I4                   9.15768482 1.6902812  5.41784680 6.032107e-08
## std98_base_I5                   11.77177934 1.3712231  8.58487505 9.093521e-18
## std98_base_I6                   10.49573115 1.4770295  7.10597259 1.194781e-12
## avgtest96                      -0.37657463 1.4765968 -0.25502874 7.987009e-01
## treatment:female_baseline     -3.98006181 2.0074746 -1.98262127 4.740975e-02

```

```

model <- lm.cluster(data = deworming,
                      formula = as.formula(paste("totyrs_enrolled ~ treatment * BMI +", controls)),
                      cluster = "psdpsch98",
                      weights = deworming$weight)
summary(model)

## R^2= 0.29369
##
##                               Estimate Std. Error      t value      Pr(>|t|)
## (Intercept)            9.3646592209 0.234111260 40.0008920 0.000000e+00
## treatment              0.4438721085 0.174112229  2.5493448 1.079255e-02
## BMI                  -0.0007735977 0.002551523 -0.3031906 7.617446e-01
## saturation_dm          1.1453317833 0.836062926  1.3699110 1.707147e-01
## demeaned_popT_6k        0.1144899233 0.045023839  2.5428734 1.099450e-02
## zoneidI2             -0.3635414387 0.295246590 -1.2313146 2.182052e-01
## zoneidI3             -0.3458168260 0.267003321 -1.2951780 1.952588e-01
## zoneidI4             -0.7307262629 0.240030062 -3.0443114 2.332136e-03
## zoneidI5             -0.3450610336 0.189399671 -1.8218671 6.847515e-02
## zoneidI6             -0.3375734980 0.252079959 -1.3391525 1.805210e-01
## zoneidI7             -0.2061321001 0.223732079 -0.9213346 3.568758e-01
## zoneidI8             -0.3090969983 0.203399064 -1.5196579 1.285970e-01
## pup_pop                0.0557231043 0.314723774  0.1770540 8.594660e-01
## month_interviewI2     -0.0523600501 0.206251381 -0.2538652 7.995997e-01
## month_interviewI3     -0.2621192547 0.248998591 -1.0526937 2.924814e-01
## month_interviewI4      0.0757273398 0.221913901  0.3412465 7.329180e-01
## month_interviewI5      0.0927886462 0.180002940  0.5154841 6.062148e-01
## month_interviewI6     -0.5180237170 0.210853129 -2.4567988 1.401812e-02
## month_interviewI7     -0.1132420960 0.218377555 -0.5185611 6.040669e-01
## month_interviewI8     -0.1064480454 0.195107800 -0.5455858 5.853507e-01
## month_interviewI9     -0.5581716691 0.198566523 -2.8110059 4.938688e-03
## month_interviewI10    -0.4078057205 0.214195121 -1.9038983 5.692343e-02
## month_interviewI11    0.0255509230 0.218319882  0.1170343 9.068328e-01
## month_interviewI12    0.2198688323 0.216728291  1.0144907 3.103487e-01
## cost_sharing           -0.2574411821 0.138012149 -1.8653516 6.213214e-02
## std98_base_I2          -0.8292292832 0.136242165 -6.0864365 1.154514e-09
## std98_base_I3          -1.6326722055 0.133233180 -12.2542463 1.594247e-34
## std98_base_I4          -2.4163845702 0.149874990 -16.1226671 1.768108e-58
## std98_base_I5          -3.5296716325 0.149498761 -23.6100394 3.039633e-123
## std98_base_I6          -4.2121406691 0.171727176 -24.5280961 7.409797e-133
## female_baseline        -0.8336787999 0.102419222 -8.1398665 3.957138e-16
## avgtest96                 0.6321743633 0.226136671  2.7955411 5.181291e-03
## treatment:BMI          -0.0063678821 0.004933846 -1.2906528 1.968241e-01

model <- lm.cluster(data = deworming,
                      formula = as.formula(paste("passed_primary_exam ~ treatment * BMI +", controls)),
                      cluster = "psdpsch98",
                      weights = deworming$weight)
summary(model)

## R^2= 0.07103
##
##                               Estimate Std. Error      t value      Pr(>|t|)
## (Intercept)            5.535869e-01 0.0428745094 12.9117963 3.861856e-38
## treatment              6.547062e-02 0.0343669594  1.9050453 5.677418e-02

```

```

## BMI          5.031535e-05 0.0004835742  0.1040489 9.171306e-01
## saturation_dm 2.361949e-01 0.1615823533  1.4617616 1.438066e-01
## demeaned_popT_6k 6.389072e-03 0.0096340051  0.6631793 5.072157e-01
## zoneidI2      -2.099851e-02 0.0539227766 -0.3894183 6.969668e-01
## zoneidI3      -3.204086e-02 0.0564256804 -0.5678418 5.701424e-01
## zoneidI4      -6.631414e-02 0.0420056195 -1.5786969 1.144056e-01
## zoneidI5      -4.800881e-02 0.0415380824 -1.1557781 2.477719e-01
## zoneidI6      -9.572677e-02 0.0441460903 -2.1684089 3.012759e-02
## zoneidI7      -1.657817e-02 0.0394938621 -0.4197657 6.746566e-01
## zoneidI8      -3.853736e-02 0.0359104956 -1.0731503 2.832037e-01
## pup_pop        -4.514183e-02 0.0578363234 -0.7805099 4.350908e-01
## month_interviewI2 -2.201490e-02 0.0372344513 -0.5912508 5.543524e-01
## month_interviewI3 2.111711e-02 0.0462636202  0.4564518 6.480651e-01
## month_interviewI4 2.681867e-02 0.0455560624  0.5886960 5.560652e-01
## month_interviewI5 8.556406e-02 0.0349825970  2.4459037 1.444896e-02
## month_interviewI6 -3.070777e-02 0.0356320416 -0.8618021 3.887964e-01
## month_interviewI7 6.276595e-02 0.0365435680  1.7175650 8.587599e-02
## month_interviewI8 1.756060e-02 0.0337730375  0.5199590 6.030921e-01
## month_interviewI9 -6.421549e-02 0.03353336462 -1.9149571 5.549798e-02
## month_interviewI10 -1.514748e-02 0.0370899935 -0.4083980 6.829815e-01
## month_interviewI11 3.718342e-02 0.0375911577  0.9891533 3.225881e-01
## month_interviewI12 9.027188e-02 0.0445472985  2.0264276 4.272098e-02
## cost_sharing    -3.986747e-02 0.0268342882 -1.4856914 1.373608e-01
## std98_base_I2   4.298428e-02 0.0303574649  1.4159378 1.567937e-01
## std98_base_I3   1.003575e-01 0.0344601827  2.9122749 3.588068e-03
## std98_base_I4   9.113439e-02 0.0274968453  3.3143579 9.185386e-04
## std98_base_I5   9.628811e-02 0.0313493095  3.0714586 2.130157e-03
## std98_base_I6   9.136947e-02 0.0265130797  3.4462038 5.685215e-04
## female_baseline -1.818531e-01 0.0191561749 -9.4931854 2.240799e-21
## avgtest96       1.396096e-01 0.0389484214  3.5844741 3.377581e-04
## treatment:BMI   -6.233804e-04 0.0006214238 -1.0031485 3.157892e-01

model <- lm.cluster(data = deworming,
                      formula = as.formula(paste("total_hours ~ treatment * BMI +", controls)),
                      cluster = "psdpsch98",
                      weights = deworming$weight)
summary(model)

## R^2= 0.0601
##
##                               Estimate Std. Error     t value   Pr(>|t|) 
## (Intercept)            13.50525307 2.39254986  5.64471124 1.654587e-08
## treatment             -0.08807024 1.32216985 -0.06661038 9.468919e-01
## BMI                  -0.01185702 0.00795430 -1.49064276 1.360553e-01
## saturation_dm         9.14797537 7.82400692  1.16921872 2.423155e-01
## demeaned_popT_6k       0.17865632 0.36507337  0.48937101 6.245791e-01
## zoneidI2              1.19198206 2.30905613  0.51622048 6.057004e-01
## zoneidI3              0.43577971 2.32919981  0.18709417 8.515868e-01
## zoneidI4              2.77665925 2.24488061  1.23688504 2.161298e-01
## zoneidI5              1.56571010 1.99267220  0.78573390 4.320234e-01
## zoneidI6              3.15765232 1.90213389  1.66005786 9.690281e-02
## zoneidI7              -0.42206929 2.01006667 -0.20997776 8.336850e-01
## zoneidI8              1.18849228 1.93314914  0.61479596 5.386895e-01
## pup_pop                -2.04011136 2.92220372 -0.69814139 4.850888e-01
## month_interviewI2      0.49695193 2.36151385  0.21043786 8.333259e-01

```

```

## month_interviewI3 -0.07874510 1.70394118 -0.04621351 9.631401e-01
## month_interviewI4  2.32142803 2.18220024  1.06380156 2.874186e-01
## month_interviewI5  0.43092732 1.89169285  0.22779984 8.198019e-01
## month_interviewI6 -0.18308257 1.80455823 -0.10145562 9.191888e-01
## month_interviewI7  0.33091493 1.86673386  0.17726947 8.592967e-01
## month_interviewI8  2.69541204 1.65496084  1.62868630 1.033794e-01
## month_interviewI9  3.34520747 1.76722218  1.89291845 5.836872e-02
## month_interviewI10 4.28160344 1.99470021  2.14648969 3.183393e-02
## month_interviewI11 2.16612547 1.66627303  1.29998232 1.936070e-01
## month_interviewI12 2.36694311 2.30869590  1.02522949 3.052549e-01
## cost_sharing      -1.64424476 0.84899263 -1.93670086 5.278192e-02
## std98_base_I2     2.77568930 1.19497558  2.32280002 2.018990e-02
## std98_base_I3     5.22062567 1.41278058  3.69528414 2.196410e-04
## std98_base_I4     9.12371865 1.72074782  5.30218231 1.144265e-07
## std98_base_I5    11.65337805 1.36540937  8.53471368 1.405029e-17
## std98_base_I6    10.35197157 1.48655608  6.96372759 3.313851e-12
## female_baseline   -6.70677094 0.97131436 -6.90484068 5.025980e-12
## avgtest96        -0.51408388 1.51685141 -0.33891512 7.346737e-01
## treatment:BMI    0.07479562 0.03886328  1.92458342 5.428150e-02

```