ECMA 31360 PSet 3

Phalgun Garimella

Part 1: Describe the Data

Question 1

```
control = read.csv("nswre74_control.csv")
treated = read.csv("nswre74_treated.csv")
df = rbind(control, treated)
dplyr::tally(dplyr::group_by(df, treat))
```

```
## # A tibble: 2 x 2
##   treat     n
##   <int> <int>
## 1     0   260
## 2     1   185
```

```
dplyr::group_by(df, treat) %>% dplyr::summarise_all(list(mean))
```

```
## # A tibble: 2 x 12
##   treat   age   edu black   hisp married nodegree  re74  re75  re78
##   <int> <dbl> <dbl> <dbl>  <dbl>   <dbl>    <dbl> <dbl> <dbl> <dbl>
## 1     0  25.1  10.1 0.827 0.108    0.154    0.835 2107. 1267. 4555.
## 2     1  25.8  10.3 0.843 0.0595   0.189    0.708 2096. 1532. 6349.
## # ... with 2 more variables: u74 <dbl>, u75 <dbl>
```

We can insert table here/on shared Overleaf document.

Part 2: Test Balance

Question 1

```
columns = colnames(df)
opvs = columns[-c(1,10)]
p_vals = c()
formula_list = c()
```

```r
null_system = c() # for Q2.2b
intercepts = c() # for Q2.2b
coefficients = c() # for Q2.2b
for (i in opvs) {
  formula = formula(paste(i, " ~ ", columns[1]))
  null_formula = formula(paste(i, " ~ ", 1))
  formula_list = c(formula_list, formula)
  null_system = c(null_system, null_formula)
  lm_model = lm(formula, data = df)
  summary_coeffs = summary(lm_model)$coefficients
  print(i)
  print(summary_coeffs)
  p_vals = c(p_vals, summary_coeffs[2,4])
  intercepts = c(intercepts, summary_coeffs[1,1])
  coefficients = c(coefficients, summary_coeffs[2,1])
}
```

```
## [1] "age"
##             Estimate Std. Error  t value     Pr(>|t|)
## (Intercept) 25.05385   0.440218 56.91230 7.88501e-206
## treat        0.76237   0.682751  1.11661  2.64764e-01
## [1] "edu"
##              Estimate Std. Error  t value     Pr(>|t|)
## (Intercept) 10.088462   0.110988 90.89690 1.27848e-288
## treat        0.257484   0.172135  1.49583  1.35411e-01
## [1] "black"
##              Estimate Std. Error   t value     Pr(>|t|)
## (Intercept) 0.8269231  0.0231384 35.738160 1.37912e-132
## treat       0.0163202  0.0358862  0.454776  6.49493e-01
## [1] "hisp"
##               Estimate Std. Error  t value        Pr(>|t|)
## (Intercept)  0.1076923  0.0175141  6.14891 0.0000000017434
## treat       -0.0482328  0.0271632 -1.77567 0.0764738932509
## [1] "married"
##             Estimate Std. Error  t value          Pr(>|t|)
## (Intercept) 0.153846  0.0232430 6.619030 0.000000000104714
## treat       0.035343  0.0360484 0.980432 0.327408104868788
## [1] "nodegree"
##              Estimate Std. Error  t value     Pr(>|t|)
## (Intercept)  0.834615  0.0253687 32.89946 5.16574e-121
## treat       -0.126507  0.0393452 -3.21532  1.39835e-03
## [1] "re74"
##             Estimate Std. Error    t value          Pr(>|t|)
## (Intercept) 2107.027    333.010  6.3272117 0.000000000611661
## treat        -11.453    516.478 -0.0221751 0.982318253353549
## [1] "re75"
##             Estimate Std. Error  t value          Pr(>|t|)
## (Intercept) 1266.909    195.466 6.481479 0.000000000242483
## treat        265.146    303.155 0.874621 0.382253831480420
```

```
## [1] "u74"
##               Estimate Std. Error   t value    Pr(>|t|)
## (Intercept)  0.7500000  0.0274815 27.291068 6.33654e-97
## treat       -0.0418919  0.0426221 -0.982868 3.26209e-01
## [1] "u75"
##               Estimate Std. Error  t value    Pr(>|t|)
## (Intercept)  0.6846154  0.0295446 23.17230 2.16054e-78
## treat       -0.0846154  0.0458218 -1.84662 6.54690e-02
```

In order to make conclusions regarding balance of OPVs, let us output p-values for the 10 t-tests below for convenience:

```
out = t(matrix(round(p_vals, 4)))
colnames(out) = opvs
rownames(out) = "p-values"
out
```

```
##             age    edu  black   hisp married nodegree   re74   re75
## p-values 0.2648 0.1354 0.6495 0.0765  0.3274   0.0014 0.9823 0.3823
##             u74    u75
## p-values 0.3262 0.0655
```

We observe that at the 5% level, we reject the null hypothesis that the mean for `nodegree` is the same in the control and treated groups. For the 9 other OPVs, however, we fail to reject the null hypothesis that each variable's mean is the same in the control and treated groups.

Question 2

a

```
sur_fit = systemfit::systemfit(formula_list, data = df, method = "SUR")
summary(sur_fit)
```

```
##
## systemfit results
## method: SUR
##
##            N   DF          SSR     detRCov  OLS-R2 McElroy-R2
## system 4450 4430 17173687963 15749397122 0.00044     0.0046
##
##        N  DF             SSR           MSE       RMSE      R2
## eq1  445 443       22320.998       50.38600    7.09831 0.00281
## eq2  445 443        1418.825        3.20276    1.78963 0.00503
## eq3  445 443          61.666        0.13920    0.37310 0.00047
## eq4  445 443          35.331        0.07975    0.28241 0.00707
## eq5  445 443          62.225        0.14046    0.37478 0.00217
```

```
## eq6  445 443          74.126        0.16733    0.40906 0.02280
## eq7  445 443 12772984837.660 28832922.88411 5369.62968 0.00000
## eq8  445 443  4400678965.093  9933812.56229 3151.79513 0.00172
## eq9  445 443          86.988        0.19636    0.44313 0.00218
## eq10 445 443         100.538        0.22695    0.47639 0.00764
##        Adj R2
## eq1   0.00056
## eq2   0.00278
## eq3  -0.00179
## eq4   0.00483
## eq5  -0.00009
## eq6   0.02060
## eq7  -0.00226
## eq8  -0.00053
## eq9  -0.00008
## eq10  0.00540
##
## The covariance matrix of the residuals used for estimation
##             eq1        eq2         eq3          eq4         eq5
## eq1    50.38600   0.244974   0.2282017  -0.17265479   0.5536763
## eq2     0.24497   3.202765   0.0293752  -0.07061474   0.0527165
## eq3     0.22820   0.029375   0.1392000  -0.07320434   0.0031819
## eq4    -0.17265  -0.070615  -0.0732043   0.07975296   0.0013797
## eq5     0.55368   0.052717   0.0031819   0.00137974   0.1404617
## eq6    -0.30837  -0.463917   0.0069816   0.00867174  -0.0048948
## eq7   -44.77087 397.532246  10.9580485 -43.48312001 287.2264039
## eq8  1130.71668 129.673476 -67.4371176  39.60097283 314.7993099
## eq9     0.33087  -0.071262   0.0051583  -0.00629614  -0.0130559
## eq10    0.27272  -0.104167   0.0072408   0.00052092  -0.0166696
##               eq6          eq7          eq8           eq9
## eq1    -0.3083723      -44.771     1130.717     0.3308706
## eq2    -0.4639166      397.532      129.673    -0.0712617
## eq3     0.0069816       10.958      -67.437     0.0051583
## eq4     0.0086717      -43.483       39.601    -0.0062961
## eq5    -0.0048948      287.226      314.799    -0.0130559
## eq6     0.1673280     -151.065       35.318     0.0078732
## eq7  -151.0653318 28832922.884 11098914.590 -1547.1565481
## eq8    35.3184906 11098914.590  9933812.562  -787.3606508
## eq9     0.0078732    -1547.157     -787.361     0.1963608
## eq10    0.0131794    -1242.491     -892.930     0.1532731
##              eq10
## eq1      0.27272096
## eq2     -0.10416739
## eq3      0.00724084
## eq4      0.00052092
## eq5     -0.01666956
## eq6      0.01317937
## eq7  -1242.49078822
## eq8   -892.92989693
## eq9      0.15327314
```

```
## eq10     0.22694912
##
## The covariance matrix of the residuals
##             eq1        eq2         eq3          eq4         eq5
## eq1    50.38600   0.244974   0.2282017  -0.17265479   0.5536763
## eq2     0.24497   3.202765   0.0293752  -0.07061474   0.0527165
## eq3     0.22820   0.029375   0.1392000  -0.07320434   0.0031819
## eq4    -0.17265  -0.070615  -0.0732043   0.07975296   0.0013797
## eq5     0.55368   0.052717   0.0031819   0.00137974   0.1404617
## eq6    -0.30837  -0.463917   0.0069816   0.00867174  -0.0048948
## eq7   -44.77087 397.532246  10.9580485 -43.48312001 287.2264039
## eq8  1130.71668 129.673476 -67.4371176  39.60097283 314.7993099
## eq9     0.33087  -0.071262   0.0051583  -0.00629614  -0.0130559
## eq10    0.27272  -0.104167   0.0072408   0.00052092  -0.0166696
##               eq6          eq7          eq8           eq9
## eq1    -0.3083723      -44.771     1130.717     0.3308706
## eq2    -0.4639166      397.532      129.673    -0.0712617
## eq3     0.0069816       10.958      -67.437     0.0051583
## eq4     0.0086717      -43.483       39.601    -0.0062961
## eq5    -0.0048948      287.226      314.799    -0.0130559
## eq6     0.1673280     -151.065       35.318     0.0078732
## eq7  -151.0653318 28832922.884 11098914.590 -1547.1565481
## eq8    35.3184906 11098914.590  9933812.562  -787.3606508
## eq9     0.0078732    -1547.157     -787.361     0.1963608
## eq10    0.0131794    -1242.491     -892.930     0.1532731
##               eq10
## eq1      0.27272096
## eq2     -0.10416739
## eq3      0.00724084
## eq4      0.00052092
## eq5     -0.01666956
## eq6      0.01317937
## eq7  -1242.49078822
## eq8   -892.92989693
## eq9      0.15327314
## eq10     0.22694912
##
## The correlations of the residuals
##             eq1       eq2        eq3       eq4       eq5       eq6
## eq1   1.0000000  0.019284  0.0861677 -0.086129  0.208124 -0.106203
## eq2   0.0192843  1.000000  0.0439945 -0.139720  0.078597 -0.633714
## eq3   0.0861677  0.043994  1.0000000 -0.694774  0.022755  0.045745
## eq4  -0.0861292 -0.139720 -0.6947744  1.000000  0.013036  0.075067
## eq5   0.2081239  0.078597  0.0227553  0.013036  1.000000 -0.031928
## eq6  -0.1062028 -0.633714  0.0457455  0.075067 -0.031928  1.000000
## eq7  -0.0011746  0.041368  0.0054698 -0.028675  0.142725 -0.068776
## eq8   0.0505407  0.022990 -0.0573484  0.044491  0.266500  0.027394
## eq9   0.1051903 -0.089860  0.0312005 -0.050312 -0.078614  0.043435
## eq10  0.0806491 -0.122181  0.0407385  0.003872 -0.093364  0.067631
##             eq7       eq8       eq9      eq10
```

```
## eq1  -0.0011746  0.050541  0.105190  0.080649
## eq2   0.0413681  0.022990 -0.089860 -0.122181
## eq3   0.0054698 -0.057348  0.031200  0.040739
## eq4  -0.0286750  0.044491 -0.050312  0.003872
## eq5   0.1427253  0.266500 -0.078614 -0.093364
## eq6  -0.0687759  0.027394  0.043435  0.067631
## eq7   1.0000000  0.655810 -0.650223 -0.485718
## eq8   0.6558102  1.000000 -0.563752 -0.594696
## eq9  -0.6502233 -0.563752  1.000000  0.726063
## eq10 -0.4857184 -0.594696  0.726063  1.000000
##
##
## SUR estimates for 'eq1' (equation 1)
## Model Formula: age ~ treat
##
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 25.05385    0.44022 56.9123   <2e-16 ***
## treat        0.76237    0.68275  1.1166   0.2648
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.09831 on 443 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 443
## SSR: 22320.99751 MSE: 50.386 Root MSE: 7.09831
## Multiple R-Squared: 0.00281 Adjusted R-Squared: 0.00056
##
##
## SUR estimates for 'eq2' (equation 2)
## Model Formula: edu ~ treat
##
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 10.08846    0.11099 90.8969   <2e-16 ***
## treat        0.25748    0.17214  1.4958   0.1354
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.78963 on 443 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 443
## SSR: 1418.82484 MSE: 3.20276 Root MSE: 1.78963
## Multiple R-Squared: 0.00503 Adjusted R-Squared: 0.00278
##
##
## SUR estimates for 'eq3' (equation 3)
## Model Formula: black ~ treat
##
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.826923   0.023138 35.7382   <2e-16 ***
## treat       0.016320   0.035886  0.4548   0.6495
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.3731 on 443 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 443
## SSR: 61.66559 MSE: 0.1392 Root MSE: 0.3731
## Multiple R-Squared: 0.00047 Adjusted R-Squared: -0.00179
##
##
## SUR estimates for 'eq4' (equation 4)
## Model Formula: hisp ~ treat
##
##              Estimate Std. Error t value       Pr(>|t|)
## (Intercept)  0.107692   0.017514  6.1489 0.000000001743 ***
## treat       -0.048233   0.027163 -1.7757        0.07647 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.28241 on 443 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 443
## SSR: 35.33056 MSE: 0.07975 Root MSE: 0.28241
## Multiple R-Squared: 0.00707 Adjusted R-Squared: 0.00483
##
##
## SUR estimates for 'eq5' (equation 5)
## Model Formula: married ~ treat
##
##             Estimate Std. Error t value        Pr(>|t|)
## (Intercept) 0.153846   0.023243  6.6190 0.0000000001047 ***
## treat       0.035343   0.036048  0.9804          0.3274
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.37478 on 443 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 443
## SSR: 62.22453 MSE: 0.14046 Root MSE: 0.37478
## Multiple R-Squared: 0.00217 Adjusted R-Squared: -0.00009
##
##
## SUR estimates for 'eq6' (equation 6)
## Model Formula: nodegree ~ treat
##
##              Estimate Std. Error t value  Pr(>|t|)
## (Intercept)  0.834615   0.025369 32.8995 < 2.2e-16 ***
## treat       -0.126507   0.039345 -3.2153  0.001398 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.40906 on 443 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 443
## SSR: 74.1263 MSE: 0.16733 Root MSE: 0.40906
## Multiple R-Squared: 0.0228 Adjusted R-Squared: 0.0206
```

```
##
##
## SUR estimates for 'eq7' (equation 7)
## Model Formula: re74 ~ treat
##
##             Estimate Std. Error t value        Pr(>|t|)
## (Intercept) 2107.027    333.010  6.3272 0.0000000006117 ***
## treat        -11.453    516.478 -0.0222          0.9823
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5369.62968 on 443 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 443
## SSR: 12772984837.6604 MSE: 28832922.88411 Root MSE: 5369.62968
## Multiple R-Squared: 0 Adjusted R-Squared: -0.00226
##
##
## SUR estimates for 'eq8' (equation 8)
## Model Formula: re75 ~ treat
##
##             Estimate Std. Error t value        Pr(>|t|)
## (Intercept)  1266.91     195.47  6.4815 0.0000000002425 ***
## treat         265.15     303.16  0.8746          0.3823
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3151.79513 on 443 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 443
## SSR: 4400678965.09339 MSE: 9933812.56229 Root MSE: 3151.79513
## Multiple R-Squared: 0.00172 Adjusted R-Squared: -0.00053
##
##
## SUR estimates for 'eq9' (equation 9)
## Model Formula: u74 ~ treat
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.750000   0.027482 27.2911   <2e-16 ***
## treat       -0.041892   0.042622 -0.9829   0.3262
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.44313 on 443 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 443
## SSR: 86.98784 MSE: 0.19636 Root MSE: 0.44313
## Multiple R-Squared: 0.00218 Adjusted R-Squared: -0.00008
##
##
## SUR estimates for 'eq10' (equation 10)
## Model Formula: u75 ~ treat
##
```

```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.684615   0.029545 23.1723  < 2e-16 ***
## treat       -0.084615   0.045822 -1.8466  0.06547 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.47639 on 443 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 443
## SSR: 100.53846 MSE: 0.22695 Root MSE: 0.47639
## Multiple R-Squared: 0.00764 Adjusted R-Squared: 0.0054
```

There does not appear to be a difference between the estimated coefficients and their SEs using the SUR system and those obtained in **Question 1**.

b

```
null_fit = systemfit(null_system, data = df, method = "SUR")
summary(null_fit)
```

```
##
## systemfit results
## method: SUR
##
##           N   DF         SSR     detRCov OLS-R2 McElroy-R2
## system 4450 4440 17181301210 16109620055      0          0
##
##        N  DF             SSR            MSE       RMSE R2 Adj R2
## eq1  445 444       22383.820       50.41401    7.10028  0      0
## eq2  445 444        1425.991        3.21169    1.79212  0      0
## eq3  445 444          61.694        0.13895    0.37276  0      0
## eq4  445 444          35.582        0.08014    0.28309  0      0
## eq5  445 444          62.360        0.14045    0.37477  0      0
## eq6  445 444          75.856        0.17085    0.41334  0      0
## eq7  445 444 12772999015.837 28768015.80144 5363.58237  0      0
## eq8  445 444  4408277960.864  9928553.96591 3150.96080  0      0
## eq9  445 444          87.178        0.19635    0.44311  0      0
## eq10 445 444         101.312        0.22818    0.47768  0      0
##
## The covariance matrix of the residuals used for estimation
##             eq1        eq2         eq3          eq4          eq5
## eq1    50.41401   0.292211   0.2307167  -0.18121773   0.55898876
## eq2     0.29221   3.211691   0.0303320  -0.07347910   0.05481324
## eq3     0.23072   0.030332   0.1389513  -0.07323110   0.00331511
## eq4    -0.18122  -0.073479  -0.0732311   0.08013969   0.00096164
## eq5     0.55899   0.054813   0.0033151   0.00096164   0.14044944
## eq6    -0.33116  -0.470802   0.0064632   0.01013767  -0.00597226
## eq7   -46.79566 395.918992  10.8878648 -43.25070351 286.48095505
```

```
## eq8  1177.38003 146.001708 -66.2317863  36.39841263 316.37165032
## eq9     0.32235  -0.073727   0.0049803  -0.00579006  -0.01338698
## eq10    0.25640  -0.109237   0.0068883   0.00151331  -0.01736006
##               eq6          eq7          eq8          eq9
## eq1    -0.3311570      -46.796     1177.380     0.3223504
## eq2    -0.4708017      395.919      146.002    -0.0737271
## eq3     0.0064632       10.888      -66.232     0.0049803
## eq4     0.0101377      -43.251       36.398    -0.0057901
## eq5    -0.0059723      286.481      316.372    -0.0133870
## eq6     0.1708473     -150.372       27.073     0.0091457
## eq7  -150.3723703 28768015.801 11073177.761 -1543.5551595
## eq8    27.0730611 11073177.761  9928553.966  -788.2913842
## eq9     0.0091457    -1543.555     -788.291     0.1963458
## eq10    0.0157556    -1239.456     -896.381     0.1537909
##               eq10
## eq1      0.2564025
## eq2     -0.1092368
## eq3      0.0068883
## eq4      0.0015133
## eq5     -0.0173601
## eq6      0.0157556
## eq7  -1239.4564632
## eq8   -896.3806087
## eq9      0.1537909
## eq10     0.2281810
##
## The covariance matrix of the residuals
##             eq1        eq2         eq3          eq4          eq5
## eq1    50.41401   0.292211   0.2307167  -0.18121773   0.55898876
## eq2     0.29221   3.211691   0.0303320  -0.07347910   0.05481324
## eq3     0.23072   0.030332   0.1389513  -0.07323110   0.00331511
## eq4    -0.18122  -0.073479  -0.0732311   0.08013969   0.00096164
## eq5     0.55899   0.054813   0.0033151   0.00096164   0.14044944
## eq6    -0.33116  -0.470802   0.0064632   0.01013767  -0.00597226
## eq7   -46.79566 395.918992  10.8878648 -43.25070351 286.48095505
## eq8  1177.38003 146.001708 -66.2317863  36.39841263 316.37165032
## eq9     0.32235  -0.073727   0.0049803  -0.00579006  -0.01338698
## eq10    0.25640  -0.109237   0.0068883   0.00151331  -0.01736006
##               eq6          eq7          eq8          eq9
## eq1    -0.3311570      -46.796     1177.380     0.3223504
## eq2    -0.4708017      395.919      146.002    -0.0737271
## eq3     0.0064632       10.888      -66.232     0.0049803
## eq4     0.0101377      -43.251       36.398    -0.0057901
## eq5    -0.0059723      286.481      316.372    -0.0133870
## eq6     0.1708473     -150.372       27.073     0.0091457
## eq7  -150.3723703 28768015.801 11073177.761 -1543.5551595
## eq8    27.0730611 11073177.761  9928553.966  -788.2913842
## eq9     0.0091457    -1543.555     -788.291     0.1963458
## eq10    0.0157556    -1239.456     -896.381     0.1537909
##               eq10
```

```
## eq1      0.2564025
## eq2     -0.1092368
## eq3      0.0068883
## eq4      0.0015133
## eq5     -0.0173601
## eq6      0.0157556
## eq7  -1239.4564632
## eq8   -896.3806087
## eq9      0.1537909
## eq10     0.2281810
##
## The correlations of the residuals
##              eq1       eq2        eq3        eq4        eq5       eq6
## eq1   1.0000000  0.022964  0.0871710 -0.0901574  0.2100717 -0.112838
## eq2   0.0229643  1.000000  0.0454049 -0.1448348  0.0816129 -0.635575
## eq3   0.0871710  0.045405  1.0000000 -0.6939695  0.0237305  0.041948
## eq4  -0.0901574 -0.144835 -0.6939695  1.0000000  0.0090641  0.086638
## eq5   0.2100717  0.081613  0.0237305  0.0090641  1.0000000 -0.038554
## eq6  -0.1128377 -0.635575  0.0419481  0.0866383 -0.0385545  1.000000
## eq7  -0.0012288  0.041189  0.0054457 -0.0284849  0.1425216 -0.067828
## eq8   0.0526257  0.025855 -0.0563887  0.0408052  0.2679135  0.020787
## eq9   0.1024571 -0.092843  0.0301516 -0.0461582 -0.0806143  0.049934
## eq10  0.0755974 -0.127603  0.0386851  0.0111909 -0.0969731  0.079798
##              eq7       eq8       eq9      eq10
## eq1  -0.0012288  0.052626  0.102457  0.075597
## eq2   0.0411893  0.025855 -0.092843 -0.127603
## eq3   0.0054457 -0.056389  0.030152  0.038685
## eq4  -0.0284849  0.040805 -0.046158  0.011191
## eq5   0.1425216  0.267913 -0.080614 -0.096973
## eq6  -0.0678280  0.020787  0.049934  0.079798
## eq7   1.0000000  0.655201 -0.649466 -0.483767
## eq8   0.6552006  1.000000 -0.564590 -0.595538
## eq9  -0.6494660 -0.564590  1.000000  0.726574
## eq10 -0.4837674 -0.595538  0.726574  1.000000
##
##
## SUR estimates for 'eq1' (equation 1)
## Model Formula: age ~ 1
##
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept) 25.37079    0.33659  75.377 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.10028 on 444 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 444
## SSR: 22383.82022 MSE: 50.41401 Root MSE: 7.10028
## Multiple R-Squared: 0 Adjusted R-Squared: 0
##
##
```

```
## SUR estimates for 'eq2' (equation 2)
## Model Formula: edu ~ 1
##
##              Estimate Std. Error t value  Pr(>|t|)
## (Intercept) 10.195506   0.084955  120.01 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.79212 on 444 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 444
## SSR: 1425.99101 MSE: 3.21169 Root MSE: 1.79212
## Multiple R-Squared: 0 Adjusted R-Squared: 0
##
##
## SUR estimates for 'eq3' (equation 3)
## Model Formula: black ~ 1
##
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept) 0.833708   0.017671  47.181 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.37276 on 444 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 444
## SSR: 61.69438 MSE: 0.13895 Root MSE: 0.37276
## Multiple R-Squared: 0 Adjusted R-Squared: 0
##
##
## SUR estimates for 'eq4' (equation 4)
## Model Formula: hisp ~ 1
##
##             Estimate Std. Error t value        Pr(>|t|)
## (Intercept)  0.08764    0.01342  6.5307 0.0000000001794 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.28309 on 444 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 444
## SSR: 35.58202 MSE: 0.08014 Root MSE: 0.28309
## Multiple R-Squared: 0 Adjusted R-Squared: 0
##
##
## SUR estimates for 'eq5' (equation 5)
## Model Formula: married ~ 1
##
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept) 0.168539   0.017766  9.4868 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 0.37477 on 444 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 444
## SSR: 62.35955 MSE: 0.14045 Root MSE: 0.37477
## Multiple R-Squared: 0 Adjusted R-Squared: 0
##
##
## SUR estimates for 'eq6' (equation 6)
## Model Formula: nodegree ~ 1
##
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept) 0.782022   0.019594  39.911 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.41334 on 444 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 444
## SSR: 75.85618 MSE: 0.17085 Root MSE: 0.41334
## Multiple R-Squared: 0 Adjusted R-Squared: 0
##
##
## SUR estimates for 'eq7' (equation 7)
## Model Formula: re74 ~ 1
##
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept)  2102.27     254.26  8.2682 1.554e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5363.58237 on 444 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 444
## SSR: 12772999015.8373 MSE: 28768015.80144 Root MSE: 5363.58237
## Multiple R-Squared: 0 Adjusted R-Squared: 0
##
##
## SUR estimates for 'eq8' (equation 8)
## Model Formula: re75 ~ 1
##
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept)  1377.14     149.37  9.2197 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3150.9608 on 444 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 444
## SSR: 4408277960.86392 MSE: 9928553.96591 Root MSE: 3150.9608
## Multiple R-Squared: 0 Adjusted R-Squared: 0
##
##
## SUR estimates for 'eq9' (equation 9)
## Model Formula: u74 ~ 1
```

```
##
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept) 0.732584   0.021005  34.876 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.44311 on 444 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 444
## SSR: 87.17753 MSE: 0.19635 Root MSE: 0.44311
## Multiple R-Squared: 0 Adjusted R-Squared: 0
##
##
## SUR estimates for 'eq10' (equation 10)
## Model Formula: u75 ~ 1
##
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept) 0.649438   0.022644   28.68 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.47768 on 444 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 444
## SSR: 101.31236 MSE: 0.22818 Root MSE: 0.47768
## Multiple R-Squared: 0 Adjusted R-Squared: 0
```

```
lrtest_obj = lrtest(null_fit, sur_fit)
lrtest_obj
```

```
## Likelihood ratio test
##
## Model 1: null_fit
## Model 2: sur_fit
##   #Df LogLik Df Chisq Pr(>Chisq)
## 1  65 -11539
## 2  75 -11529 10  20.1     0.0283 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Let us now test joint balance manually. We will be testing the following joint hypotheses:

$H_0 : $

$H_1 : $

```
n = dim(df)[1]
cov_mat = matrix(0, 10, 10)
df_opvs = df[,-c(1,10)]
for (i in 1:dim(cov_mat)[1]) {
  for (j in 1:dim(cov_mat)[2]) {
    u_i = as.matrix(df_opvs[,i] - (df[,1] * coefficients[i] + intercepts[i]))
```

```
    u_j = as.matrix(df_opvs[,j] - (df[,1] * coefficients[j] + intercepts[j]))
    cov_mat[i,j] = t(u_i) %*% u_j / (n - 2)
  }
}
cov_mat
```

```
##              [,1]        [,2]         [,3]          [,4]         [,5]
##  [1,]   50.385999   0.2449745   0.22820169  -0.172654787   0.55367627
##  [2,]    0.244974   3.2027649   0.02937517  -0.070614737   0.05271655
##  [3,]    0.228202   0.0293752   0.13919998  -0.073204338   0.00318186
##  [4,]   -0.172655  -0.0706147  -0.07320434   0.079752960   0.00137974
##  [5,]    0.553676   0.0527165   0.00318186   0.001379744   0.14046170
##  [6,]   -0.308372  -0.4639166   0.00698155   0.008671738  -0.00489481
##  [7,]  -44.770874 397.5322464  10.95804850 -43.483120007 287.22640387
##  [8,] 1130.716679 129.6734756 -67.43711762  39.600972835 314.79930993
##  [9,]    0.330871  -0.0712617   0.00515832  -0.006296138  -0.01305595
## [10,]    0.272721  -0.1041674   0.00724084   0.000520924  -0.01666956
##                [,6]          [,7]          [,8]           [,9]
##  [1,]   -0.30837232      -44.7709     1130.7167     0.33087060
##  [2,]   -0.46391664      397.5322      129.6735    -0.07126167
##  [3,]    0.00698155       10.9580      -67.4371     0.00515832
##  [4,]    0.00867174      -43.4831       39.6010    -0.00629614
##  [5,]   -0.00489481      287.2264      314.7993    -0.01305595
##  [6,]    0.16732799     -151.0653       35.3185     0.00787322
##  [7,] -151.06533176 28832922.8841 11098914.5903 -1547.15654813
##  [8,]   35.31849059 11098914.5903  9933812.5623  -787.36065078
##  [9,]    0.00787322    -1547.1565     -787.3607     0.19636081
## [10,]    0.01317937    -1242.4908     -892.9299     0.15327314
##                [,10]
##  [1,]     0.272720958
##  [2,]    -0.104167390
##  [3,]     0.007240840
##  [4,]     0.000520924
##  [5,]    -0.016669561
##  [6,]     0.013179371
##  [7,] -1242.490788222
##  [8,]  -892.929896927
##  [9,]     0.153273138
## [10,]     0.226949123
```

```
null_cov_mat = matrix(0, 10, 10)
null_coeffs = c()
for (i in opvs) {
  null_formula = formula(paste(i, " ~ ", 1))
  null_lm_model = lm(null_formula, data = df)
  null_summary_coeffs = summary(null_lm_model)$coefficients
  null_coeffs = c(null_coeffs, null_summary_coeffs[1,1])
}
```

```
for (i in 1:dim(null_cov_mat)[1]) {
  for (j in 1:dim(null_cov_mat)[2]) {
    u_i = as.matrix(df_opvs[,i] - null_coeffs[i])
    u_j = as.matrix(df_opvs[,j] - null_coeffs[j])
    null_cov_mat[i,j] = t(u_i) %*% u_j / (n - 1)
  }
}
null_cov_mat
```

```
##              [,1]        [,2]         [,3]          [,4]
##  [1,]   50.414010   0.2922108   0.23071667  -0.181217735
##  [2,]    0.292211   3.2116915   0.03033202  -0.073479097
##  [3,]    0.230717   0.0303320   0.13895131  -0.073231096
##  [4,]   -0.181218  -0.0734791  -0.07323110   0.080139690
##  [5,]    0.558989   0.0548132   0.00331511   0.000961636
##  [6,]   -0.331157  -0.4708017   0.00646320   0.010137666
##  [7,]  -46.795659 395.9189923  10.88786476 -43.250703514
##  [8,] 1177.380034 146.0017082 -66.23178630  36.398412629
##  [9,]    0.322350  -0.0737271   0.00498026  -0.005790060
## [10,]    0.256402  -0.1092368   0.00688835   0.001513311
##                [,5]          [,6]          [,7]          [,8]
##  [1,]   0.558988764   -0.33115700      -46.7957     1177.3800
##  [2,]   0.054813240   -0.47080170      395.9190      146.0017
##  [3,]   0.003315113    0.00646320       10.8879      -66.2318
##  [4,]   0.000961636    0.01013767      -43.2507       36.3984
##  [5,]   0.140449438   -0.00597226      286.4810      316.3717
##  [6,]  -0.005972264    0.17084725     -150.3724       27.0731
##  [7,] 286.480955046 -150.37237032 28768015.8014 11073177.7610
##  [8,] 316.371650323   27.07306112 11073177.7610  9928553.9659
##  [9,]  -0.013386982    0.00914566    -1543.5552     -788.2914
## [10,]  -0.017360057    0.01575564    -1239.4565     -896.3806
##                [,9]          [,10]
##  [1,]     0.32235044     0.25640247
##  [2,]    -0.07372710    -0.10923676
##  [3,]     0.00498026     0.00688835
##  [4,]    -0.00579006     0.00151331
##  [5,]    -0.01338698    -0.01736006
##  [6,]     0.00914566     0.01575564
##  [7,] -1543.55515945 -1239.45646319
##  [8,]  -788.29138417  -896.38060869
##  [9,]     0.19634578     0.15379087
## [10,]     0.15379087     0.22818099
```

```
lr_test_stat = (n - 2) * (log(norm(null_cov_mat)) - log(norm(cov_mat)))
chi_sq_test_stat = -2 * lr_test_stat * 10
chi_sq_test_stat
```

```
## [1] 20.1347
```

```r
pchisq(chi_sq_test_stat, df = 10, lower.tail = FALSE)
```

```
## [1] 0.0280045
```