ECMA 31360, PSet 2: Solutions

YOUR NAMES, University of Chicago

XX-XX-2026

Carry-over from PSet2

([ ] out of 50p) PART I: Test Balance in the Observed Predetermined Variables (OPVs)

([ ] out of 15p) Q1: Read and Understand the NSW Application's Companion Document

Done!

([ ] out of 23p) Q2: Implement Procedure 4 (SUR Estimation followed by Joint Testing)

([ ] out of 10p) Q2.a: SUR Estimation

```r
# Load packages
library(systemfit)
```

```
## Warning: package 'systemfit' was built under R version 4.5.1

## Loading required package: Matrix

## Loading required package: car

## Warning: package 'car' was built under R version 4.5.2

## Loading required package: carData

## Warning: package 'carData' was built under R version 4.5.1

## Loading required package: lmtest

## Warning: package 'lmtest' was built under R version 4.5.1

## Loading required package: zoo

## Warning: package 'zoo' was built under R version 4.5.1
```

```
##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

##
## Please cite the 'systemfit' package as:
## Arne Henningsen and Jeff D. Hamann (2007). systemfit: A Package for Estimating Systems of Simultaneous Equati
##
## If you have questions, suggestions, or comments regarding the 'systemfit' package, please use a forum or 'tra
## https://r-forge.r-project.org/projects/systemfit/
```

```r
# Load data and create treatment indicator
treated <- read.csv("nswre74_treated.csv")
control <- read.csv("nswre74_control.csv")
treated$treat <- 1
control$treat <- 0
df <- rbind(treated, control)

# OPVs
opvs <- c("age","edu","nodegree","black","hisp",
          "married","u74","u75","re74","re75")

# SUR system: one equation per OPV
sur_system <- setNames(
  lapply(opvs, function(v) as.formula(paste(v, "~ treat"))),
  opvs
)

# Estimate SUR (FGLS)
sur_fit <- systemfit(sur_system, data = df, method = "SUR")

summary(sur_fit)
```

```
##
## systemfit results
## method: SUR
##
##           N   DF         SSR     detRCov   OLS-R2 McElroy-R2
## system 4450 4430 17173687963 15749397122 0.000443   0.004599
##
##            N  DF         SSR         MSE        RMSE       R2    Adj R2
## age      445 443 2.23210e+04 5.03860e+01    7.098310 0.002807  0.000556
## edu      445 443 1.41882e+03 3.20276e+00    1.789627 0.005025  0.002779
## nodegree 445 443 7.41263e+01 1.67328e-01    0.409057 0.022805  0.020599
## black    445 443 6.16656e+01 1.39200e-01    0.373095 0.000467 -0.001790
## hisp     445 443 3.53306e+01 7.97530e-02    0.282406 0.007067  0.004826
## married  445 443 6.22245e+01 1.40462e-01    0.374782 0.002165 -0.000087
## u74      445 443 8.69878e+01 1.96361e-01    0.443126 0.002176 -0.000077
## u75      445 443 1.00538e+02 2.26949e-01    0.476392 0.007639  0.005399
## re74     445 443 1.27730e+10 2.88329e+07 5369.629678 0.000001 -0.002256
## re75     445 443 4.40068e+09 9.93381e+06 3151.795133 0.001724 -0.000530
##
## The covariance matrix of the residuals used for estimation
##                  age        edu     nodegree        black         hisp
## age        50.385999  0.2449745 -3.08372e-01   0.22820169 -1.72655e-01
## edu         0.244974  3.2027649 -4.63917e-01   0.02937517 -7.06147e-02
## nodegree   -0.308372  -0.4639166  1.67328e-01   0.00698155  8.67174e-03
## black       0.228202   0.0293752  6.98155e-03   0.13919998 -7.32043e-02
## hisp       -0.172655  -0.0706147  8.67174e-03  -0.07320434  7.97530e-02
## married     0.553676   0.0527165 -4.89481e-03   0.00318186  1.37974e-03
## u74         0.330871  -0.0712617  7.87322e-03   0.00515832 -6.29614e-03
## u75         0.272721  -0.1041674  1.31794e-02   0.00724084  5.20924e-04
## re74      -44.770874 397.5322464 -1.51065e+02  10.95804850 -4.34831e+01
## re75     1130.716679 129.6734756  3.53185e+01 -67.43711762  3.96010e+01
##              married          u74          u75         re74         re75
## age        0.55367627  3.30871e-01  2.72721e-01 -4.47709e+01  1.13072e+03
## edu        0.05271655 -7.12617e-02 -1.04167e-01  3.97532e+02  1.29673e+02
## nodegree  -0.00489481  7.87322e-03  1.31794e-02 -1.51065e+02  3.53185e+01
## black      0.00318186  5.15832e-03  7.24084e-03  1.09580e+01 -6.74371e+01
## hisp       0.00137974 -6.29614e-03  5.20924e-04 -4.34831e+01  3.96010e+01
## married    0.14046170 -1.30559e-02 -1.66696e-02  2.87226e+02  3.14799e+02
## u74       -0.01305595  1.96361e-01  1.53273e-01 -1.54716e+03 -7.87361e+02
## u75       -0.01666956  1.53273e-01  2.26949e-01 -1.24249e+03 -8.92930e+02
## re74     287.22640387 -1.54716e+03 -1.24249e+03  2.88329e+07  1.10989e+07
## re75     314.79930993 -7.87361e+02 -8.92930e+02  1.10989e+07  9.93381e+06
##
## The covariance matrix of the residuals
##                  age         edu     nodegree        black         hisp
## age        50.385999   0.2449745 -3.08372e-01   0.22820169 -1.72655e-01
## edu         0.244974   3.2027649 -4.63917e-01   0.02937517 -7.06147e-02
## nodegree   -0.308372  -0.4639166  1.67328e-01   0.00698155  8.67174e-03
## black       0.228202   0.0293752  6.98155e-03   0.13919998 -7.32043e-02
## hisp       -0.172655  -0.0706147  8.67174e-03  -0.07320434  7.97530e-02
## married     0.553676   0.0527165 -4.89481e-03   0.00318186  1.37974e-03
## u74         0.330871  -0.0712617  7.87322e-03   0.00515832 -6.29614e-03
## u75         0.272721  -0.1041674  1.31794e-02   0.00724084  5.20924e-04
## re74      -44.770874 397.5322464 -1.51065e+02  10.95804850 -4.34831e+01
## re75     1130.716679 129.6734756  3.53185e+01 -67.43711762  3.96010e+01
##              married          u74          u75         re74         re75
## age        0.55367627  3.30871e-01  2.72721e-01 -4.47709e+01  1.13072e+03
## edu        0.05271655 -7.12617e-02 -1.04167e-01  3.97532e+02  1.29673e+02
## nodegree  -0.00489481  7.87322e-03  1.31794e-02 -1.51065e+02  3.53185e+01
## black      0.00318186  5.15832e-03  7.24084e-03  1.09580e+01 -6.74371e+01
## hisp       0.00137974 -6.29614e-03  5.20924e-04 -4.34831e+01  3.96010e+01
## married    0.14046170 -1.30559e-02 -1.66696e-02  2.87226e+02  3.14799e+02
## u74       -0.01305595  1.96361e-01  1.53273e-01 -1.54716e+03 -7.87361e+02
## u75       -0.01666956  1.53273e-01  2.26949e-01 -1.24249e+03 -8.92930e+02
## re74     287.22640387 -1.54716e+03 -1.24249e+03  2.88329e+07  1.10989e+07
## re75     314.79930993 -7.87361e+02 -8.92930e+02  1.10989e+07  9.93381e+06
##
## The correlations of the residuals
##                 age        edu   nodegree       black        hisp    married
## age       1.00000000  0.0192843 -0.1062028  0.08616766 -0.08612917  0.2081239
## edu       0.01928428  1.0000000 -0.6337137  0.04399449 -0.13972027  0.0785969
## nodegree -0.10620285 -0.6337137  1.0000000  0.04574547  0.07506690 -0.0319280
## black     0.08616766  0.0439945  0.0457455  1.00000000 -0.69477442  0.0227553
## hisp     -0.08612917 -0.1397203  0.0750669 -0.69477442  1.00000000  0.0130361
## married   0.20812393  0.0785969 -0.0319280  0.02275529  0.01303605  1.0000000
## u74       0.10519033 -0.0898599  0.0434351  0.03120047 -0.05031221 -0.0786144
## u75       0.08064906 -0.1221814  0.0676311  0.04073851  0.00387201 -0.0933643
## re74     -0.00117462  0.0413681 -0.0687759  0.00546977 -0.02867497  0.1427253
## re75      0.05054066  0.0229896  0.0273943 -0.05734842  0.04449124  0.2664998
##                 u74        u75        re74       re75
## age       0.1051903  0.08064906 -0.00117462  0.0505407
## edu      -0.0898599 -0.12218137  0.04136809  0.0229896
## nodegree  0.0434351  0.06763105 -0.06877588  0.0273943
## black     0.0312005  0.04073851  0.00546977 -0.0573484
## hisp     -0.0503122  0.00387201 -0.02867497  0.0444912
## married  -0.0786144 -0.09336432  0.14272533  0.2664998
## u74       1.0000000  0.72606314 -0.65022329 -0.5637523
## u75       0.7260631  1.00000000 -0.48571844 -0.5946962
## re74     -0.6502233 -0.48571844  1.00000000  0.6558102
## re75     -0.5637523 -0.59469615  0.65581024  1.0000000
##
##
## SUR estimates for 'age' (equation 1)
## Model Formula: age ~ treat
## <environment: 0x000002716988c190>
##
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept) 25.053846   0.440218 56.91230  < 2e-16 ***
## treat        0.762370   0.682751  1.11661  0.26476
## ---
## 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: 22320.997505 MSE: 50.385999 Root MSE: 7.09831
## Multiple R-Squared: 0.002807 Adjusted R-Squared: 0.000556
##
##
## SUR estimates for 'edu' (equation 2)
## Model Formula: edu ~ treat
## <environment: 0x00000271698afaf0>
##
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept) 10.088462   0.110988 90.89690  < 2e-16 ***
## treat        0.257484   0.172135  1.49583  0.13541
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.789627 on 443 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 443
## SSR: 1418.824844 MSE: 3.202765 Root MSE: 1.789627
## Multiple R-Squared: 0.005025 Adjusted R-Squared: 0.002779
##
##
## SUR estimates for 'nodegree' (equation 3)
## Model Formula: nodegree ~ treat
## <environment: 0x0000027169999b28>
##
##               Estimate Std. Error  t value   Pr(>|t|)
## (Intercept)  0.8346154  0.0253687 32.89946 < 2.22e-16 ***
## treat       -0.1265073  0.0393452 -3.21532  0.0013984 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.409057 on 443 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 443
## SSR: 74.126299 MSE: 0.167328 Root MSE: 0.409057
## Multiple R-Squared: 0.022805 Adjusted R-Squared: 0.020599
##
##
## SUR estimates for 'black' (equation 4)
## Model Formula: black ~ treat
## <environment: 0x0000027169998388>
```

```
##
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept) 0.8269231  0.0231384 35.73816  < 2e-16 ***
## treat       0.0163202  0.0358862  0.45478  0.64949
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 0.373095 on 443 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 443
## SSR: 61.665593 MSE: 0.1392 Root MSE: 0.373095
## Multiple R-Squared: 0.000467 Adjusted R-Squared: -0.00179
##
##
## SUR estimates for ’hisp’ (equation 5)
## Model Formula: hisp ~ treat
## <environment: 0x00000271699b0af8>
##
##               Estimate Std. Error  t value   Pr(>|t|)
## (Intercept)  0.1076923  0.0175141  6.14891 1.7434e-09 ***
## treat       -0.0482328  0.0271632 -1.77567   0.076474 .
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 0.282406 on 443 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 443
## SSR: 35.330561 MSE: 0.079753 Root MSE: 0.282406
## Multiple R-Squared: 0.007067 Adjusted R-Squared: 0.004826
##
##
## SUR estimates for ’married’ (equation 6)
## Model Formula: married ~ treat
## <environment: 0x00000271699bb268>
##
##              Estimate Std. Error t value   Pr(>|t|)
## (Intercept) 0.1538462  0.0232430 6.61903 1.0471e-10 ***
## treat       0.0353430  0.0360484 0.98043    0.32741
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 0.374782 on 443 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 443
## SSR: 62.224532 MSE: 0.140462 Root MSE: 0.374782
## Multiple R-Squared: 0.002165 Adjusted R-Squared: -8.7e-05
##
##
## SUR estimates for ’u74’ (equation 7)
## Model Formula: u74 ~ treat
## <environment: 0x00000271699b59d8>
##
##               Estimate Std. Error  t value Pr(>|t|)
## (Intercept)  0.7500000  0.0274815 27.29107  < 2e-16 ***
## treat       -0.0418919  0.0426221 -0.98287  0.32621
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 0.443126 on 443 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 443
## SSR: 86.987838 MSE: 0.196361 Root MSE: 0.443126
## Multiple R-Squared: 0.002176 Adjusted R-Squared: -7.7e-05
##
```

```
##
## SUR estimates for ’u75’ (equation 8)
## Model Formula: u75 ~ treat
## <environment: 0x00000271699b4238>
##
##                Estimate Std. Error  t value Pr(>|t|)
## (Intercept)  0.6846154  0.0295446 23.17230  < 2e-16 ***
## treat       -0.0846154  0.0458218 -1.84662 0.065469 .
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 0.476392 on 443 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 443
## SSR: 100.538462 MSE: 0.226949 Root MSE: 0.476392
## Multiple R-Squared: 0.007639 Adjusted R-Squared: 0.005399
##
##
## SUR estimates for ’re74’ (equation 9)
## Model Formula: re74 ~ treat
## <environment: 0x00000271699ae9a8>
##
##             Estimate Std. Error  t value   Pr(>|t|)
## (Intercept) 2107.027    333.010  6.32721 6.1166e-10 ***
## treat        -11.453    516.478 -0.02218    0.98232
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 5369.629678 on 443 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 443
## SSR: 12772984837.6604 MSE: 28832922.884109 Root MSE: 5369.629678
## Multiple R-Squared: 1e-06 Adjusted R-Squared: -0.002256
##
##
## SUR estimates for ’re75’ (equation 10)
## Model Formula: re75 ~ treat
## <environment: 0x0000027169995118>
##
##             Estimate Std. Error t value   Pr(>|t|)
## (Intercept) 1266.909    195.466 6.48148 2.4248e-10 ***
## treat        265.146    303.155 0.87462    0.38225
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 3151.795133 on 443 degrees of freedom
## Number of observations: 445 Degrees of Freedom: 443
## SSR: 4400678965.09339 MSE: 9933812.562288 Root MSE: 3151.795133
## Multiple R-Squared: 0.001724 Adjusted R-Squared: -0.00053
```

Script and Output

We estimate a 10-equation SUR system where each OPV is regressed on the treatment indicator. The intercept in each equation equals the control-group mean of that OPV, and the coefficient on treat equals the treated–control difference in means. The estimated residual covariance/correlation matrices show substantial cross-equation dependence for economically related OPVs (e.g., u74–u75, re74–re75, edu–nodegree), which motivates using SUR and is required for the joint balance test in subsequent questions.

([ ] out of 1p) **Q2.b: Compare FGLS to OLS equation-by-equation**

```
# Q2.b: OLS equation-by-equation (for comparison)
ols_fits <- lapply(opvs, function(v) lm(as.formula(paste0(v, " ~ treat")), data = df))
names(ols_fits) <- opvs

# Compare treat coefficients
sapply(ols_fits, function(m) coef(m)["treat"])
```

```
##      age.treat      edu.treat nodegree.treat    black.treat     hisp.treat
##     0.76237006     0.25748441    -0.12650728     0.01632017    -0.04823285
##  married.treat      u74.treat      u75.treat     re74.treat     re75.treat
##     0.03534304    -0.04189189    -0.08461538   -11.45295788   265.14629853
```

```
coef(sur_fit)[grep("treat", names(coef(sur_fit)))]
```

```
##      age_treat      edu_treat nodegree_treat    black_treat     hisp_treat
##     0.76237006     0.25748441    -0.12650728     0.01632017    -0.04823285
##  married_treat      u74_treat      u75_treat     re74_treat     re75_treat
##     0.03534304    -0.04189189    -0.08461538   -11.45295788   265.14629853
```

Since each equation has the same regressors (an intercept term and a common treatment indicator), the point estimates from SUR/FGLS are the same as those from OLS for each equation (i.e., the estimated mean differences between the treatment and control groups are the same). The key advantage of SUR is that it can estimate the full cross-equation variance-covariance structure, which is necessary for conducting joint balance tests across different OPVs (and may lead to different systematic standard errors for linear combinations of the coefficients).

([ ] out of 3p) **Q2.c: Take a closer look at the variance-covariance matrix**

Script and Output

```
V_hat <- vcov(sur_fit)
nm <- names(coef(sur_fit))
nm[1:4]  # should look like: age_(Intercept), age_treat, edu_(Intercept), edu_treat
```

```
## [1] "age_(Intercept)" "age_treat"       "edu_(Intercept)" "edu_treat"
```

```
i01 <- 1  # pi0,1
i11 <- 2  # pi1,1
i02 <- 3  # pi0,2
i12 <- 4  # pi1,2

# (i) should be ~ 0
V_hat[i01, i11] + V_hat[i01, i01]
```

```
## [1] 1.387779e-16
```

```
# (ii) should be ~ 0
V_hat[i01, i12] + V_hat[i01, i02]
```

```
## [1] 1.832302e-17
```

```r
# (iii) should be ~ 0
V_hat[i01, i12] - V_hat[i11, i02]
```

```
## [1] 5.334275e-17
```

We estimate, for each OPV $j = 1, \dots, 10$, the regression

$$X_{ij} = \pi_{0,j} + \pi_{1,j} D_i + u_{ij},$$

where $D_i \in \{0, 1\}$ is the treatment indicator. Because the regressor is only an intercept and a binary variable, the OLS (and hence the equation-by-equation component of the SUR estimator) admits a simple "group-mean" representation:

$$\widehat{\pi}_{0,j} = \bar{X}_{j,0}, \qquad \widehat{\pi}_{1,j} = \bar{X}_{j,1} - \bar{X}_{j,0},$$

where

$$\bar{X}_{j,0} := \frac{1}{n_0} \sum_{i:D_i=0} X_{ij}, \qquad \bar{X}_{j,1} := \frac{1}{n_1} \sum_{i:D_i=1} X_{ij}.$$

Thus, $\widehat{\pi}_{0,j}$ is the control-group mean of OPV j, and $\widehat{\pi}_{1,j}$ is the treated–control difference in means for OPV j.

Expression (1) is the top-left 4×4 block of $\widehat{\mathrm{Var}}(\widehat{\pi})$ for the parameter vector

$$(\widehat{\pi}_{0,1}, \widehat{\pi}_{1,1}, \widehat{\pi}_{0,2}, \widehat{\pi}_{1,2})'.$$

The following equalities follow directly from the above identities and the (approximate) independence between the treated and control subsamples implied by unconditional random assignment (URA).

(i) Why is $\widehat{\mathrm{Cov}}[\widehat{\pi}_{0,1}, \widehat{\pi}_{1,1}] = -\widehat{\mathrm{Var}}[\widehat{\pi}_{0,1}]$? Using $\widehat{\pi}_{0,1} = \bar{X}_{1,0}$ and $\widehat{\pi}_{1,1} = \bar{X}_{1,1} - \bar{X}_{1,0}$,

$$\widehat{\mathrm{Cov}}(\widehat{\pi}_{0,1}, \widehat{\pi}_{1,1}) = \widehat{\mathrm{Cov}}\big(\bar{X}_{1,0},\ \bar{X}_{1,1} - \bar{X}_{1,0}\big) = \widehat{\mathrm{Cov}}(\bar{X}_{1,0}, \bar{X}_{1,1}) - \widehat{\mathrm{Var}}(\bar{X}_{1,0}).$$

Under URA, $\bar{X}_{1,0}$ and $\bar{X}_{1,1}$ are based on non-overlapping subsamples and are approximately independent, hence $\widehat{\mathrm{Cov}}(\bar{X}_{1,0}, \bar{X}_{1,1}) \approx 0$. Therefore,

$$\widehat{\mathrm{Cov}}(\widehat{\pi}_{0,1}, \widehat{\pi}_{1,1}) = -\widehat{\mathrm{Var}}(\bar{X}_{1,0}) = -\widehat{\mathrm{Var}}(\widehat{\pi}_{0,1}).$$

(ii) Why is $\widehat{\mathrm{Cov}}[\widehat{\pi}_{0,1}, \widehat{\pi}_{1,2}] = -\widehat{\mathrm{Cov}}[\widehat{\pi}_{0,1}, \widehat{\pi}_{0,2}]$? Using $\widehat{\pi}_{0,1} = \bar{X}_{1,0}$ and $\widehat{\pi}_{1,2} = \bar{X}_{2,1} - \bar{X}_{2,0}$,

$$\widehat{\mathrm{Cov}}(\widehat{\pi}_{0,1}, \widehat{\pi}_{1,2}) = \widehat{\mathrm{Cov}}\big(\bar{X}_{1,0},\ \bar{X}_{2,1} - \bar{X}_{2,0}\big) = \widehat{\mathrm{Cov}}(\bar{X}_{1,0}, \bar{X}_{2,1}) - \widehat{\mathrm{Cov}}(\bar{X}_{1,0}, \bar{X}_{2,0}).$$

Again, URA implies $\widehat{\mathrm{Cov}}(\bar{X}_{1,0}, \bar{X}_{2,1}) \approx 0$, so

$$\widehat{\mathrm{Cov}}(\widehat{\pi}_{0,1}, \widehat{\pi}_{1,2}) = -\widehat{\mathrm{Cov}}(\bar{X}_{1,0}, \bar{X}_{2,0}) = -\widehat{\mathrm{Cov}}(\widehat{\pi}_{0,1}, \widehat{\pi}_{0,2}),$$

since $\widehat{\pi}_{0,2} = \bar{X}_{2,0}$.

(iii) Why is $\widehat{\mathrm{Cov}}[\widehat{\pi}_{0,1}, \widehat{\pi}_{1,2}] = \widehat{\mathrm{Cov}}[\widehat{\pi}_{1,1}, \widehat{\pi}_{0,2}]$? Compute

$$\widehat{\mathrm{Cov}}(\widehat{\pi}_{1,1}, \widehat{\pi}_{0,2}) = \widehat{\mathrm{Cov}}\big(\bar{X}_{1,1} - \bar{X}_{1,0},\ \bar{X}_{2,0}\big) = \widehat{\mathrm{Cov}}(\bar{X}_{1,1}, \bar{X}_{2,0}) - \widehat{\mathrm{Cov}}(\bar{X}_{1,0}, \bar{X}_{2,0}).$$

Under URA, $\widehat{\mathrm{Cov}}(\bar{X}_{1,1}, \bar{X}_{2,0}) \approx 0$, hence

$$\widehat{\mathrm{Cov}}(\widehat{\pi}_{1,1}, \widehat{\pi}_{0,2}) = -\widehat{\mathrm{Cov}}(\bar{X}_{1,0}, \bar{X}_{2,0}).$$

But part (ii) showed

$$\widehat{\mathrm{Cov}}(\widehat{\pi}_{0,1}, \widehat{\pi}_{1,2}) = -\widehat{\mathrm{Cov}}(\bar{X}_{1,0}, \bar{X}_{2,0}),$$

so

$$\widehat{\mathrm{Cov}}(\widehat{\pi}_{0,1}, \widehat{\pi}_{1,2}) = \widehat{\mathrm{Cov}}(\widehat{\pi}_{1,1}, \widehat{\pi}_{0,2}).$$

Commentary

Because each equation includes only an intercept and a binary treatment indicator, $\widehat{\pi}_{0,j}$ equals the control-group mean and $\widehat{\pi}_{1,j}$ equals the treated–control difference in means. Under unconditional random assignment, treated and control subsamples are (approximately) independent. These two facts jointly imply the sign and symmetry restrictions in the top-left 4×4 block of $\widehat{\mathrm{Var}}(\widehat{\pi})$, and the implied equalities are verified numerically up to machine precision.

([ ] out of 4p) Q2.d: Implement Joint Test Manually

```r
# coef and vcov from SUR
b_hat <- coef(sur_fit)
V_hat <- vcov(sur_fit)
nm <- names(b_hat)

# 1) pick the J=10 treat coefficients
treat_idx <- grep("_treat$", nm)
J <- length(treat_idx)        # should be 10
K <- length(b_hat)            # should be 20

# 2) build R matrix selecting treat coefficients: R b = (pi_1,1,...,pi_1,J)'
R <- matrix(0, nrow = J, ncol = K)
for (j in 1:J) R[j, treat_idx[j]] <- 1
r0 <- rep(0, J)

# 3) compute quadratic form Q = (R b - r)' (R V R')^{-1} (R b - r)
d <- as.vector(R %*% b_hat - r0)
Q <- as.numeric(t(d) %*% solve(R %*% V_hat %*% t(R)) %*% d)

# 4) F statistic and p-value using F_{J, Jn-K}
F_stat <- Q / J
df1 <- J
n <- 445
df2 <- J*n - K                          # should be 4430
p_F <- 1 - pf(F_stat, df1 = df1, df2 = df2)

# 5) S (Wald) statistic and p-value using Chi-square_J
S_stat <- J * F_stat    # equals Q
p_S <- 1 - pchisq(S_stat, df = J)

c(J=J, K=K, n=n, df2=df2,
  F_stat=F_stat, p_value_F=p_F,
  S_stat=S_stat, p_value_S=p_S)
```

```
##            J            K            n          df2       F_stat    p_value_F
## 1.000000e+01 2.000000e+01 4.450000e+02 4.430000e+03 2.046563e+00 2.538070e-02
##       S_stat    p_value_S
## 2.046563e+01 2.514383e-02
```

Script and Output

We test the joint null hypothesis that the coefficients on `treat` are zero in all $J = 10$ equations:

$$H_0: \ \pi_{1,1} = \pi_{1,2} = \cdots = \pi_{1,10} = 0.$$

Using the companion document's expressions for Procedure 4, we compute the F statistic and the Wald (S) statistic. With $n = 445$ observations per equation and $K = 2J = 20$ parameters in the stacked system, the denominator degrees of freedom are

$$Jn - K = 10 \times 445 - 20 = 4430.$$

Our manual calculations yield

$$F = 2.0466 \quad \text{with} \quad p\text{-value} = 0.0254 \;\; \text{using } F_{10,4430},$$

and equivalently

$$S = J \cdot F = 20.4656 \quad \text{with} \quad p\text{-value} = 0.0251 \;\; \text{using } \chi^2_{10}.$$

At the 5% level, we reject H_0, indicating that the OPVs are not jointly balanced across treatment and control.

([ ] out of 4p) Q2.e: Implement Joint Test Automatically

Script and Output

```
library(car)

# Names of coefficients
names(coef(sur_fit))
```

```
##  [1] "age_(Intercept)"      "age_treat"             "edu_(Intercept)"
##  [4] "edu_treat"            "nodegree_(Intercept)" "nodegree_treat"
##  [7] "black_(Intercept)"    "black_treat"           "hisp_(Intercept)"
## [10] "hisp_treat"           "married_(Intercept)"  "married_treat"
## [13] "u74_(Intercept)"      "u74_treat"             "u75_(Intercept)"
## [16] "u75_treat"            "re74_(Intercept)"     "re74_treat"
## [19] "re75_(Intercept)"     "re75_treat"
```

```
# Joint null: all treat coefficients equal zero
# We write one restriction per equation
lh <- linearHypothesis(
  sur_fit,
  c(
    "age_treat = 0",
    "edu_treat = 0",
    "nodegree_treat = 0",
    "black_treat = 0",
    "hisp_treat = 0",
    "married_treat = 0",
    "u74_treat = 0",
    "u75_treat = 0",
    "re74_treat = 0",
    "re75_treat = 0"
  ),
  test = "F"
)

lh
```

```
## Linear hypothesis test (F statistic of a Wald test)
##
## Hypothesis:
## age_treat = 0
## edu_treat = 0
## nodegree_treat = 0
## black_treat = 0
## hisp_treat = 0
## married_treat = 0
## u74_treat = 0
## u75_treat = 0
## re74_treat = 0
## re75_treat = 0
##
## Model 1: restricted model
## Model 2: sur_fit
##
##   Res.Df Df      F  Pr(>F)
## 1   4440
## 2   4430 10 2.0466 0.02538 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
F_auto <- lh[2, "F"]
S_auto <- 10 * F_auto

c(F_auto = F_auto, S_auto = S_auto)
```

```
##    F_auto    S_auto
##  2.046563 20.465626
```

Commentary

The automatic joint test implemented via `car::linearHypothesis()` yields an F statistic of 2.0466 with a p-value of 0.02538. These values coincide exactly with those obtained from the manual construction of the Wald test in Q2.d, as both procedures rely on the same linear restrictions and the same estimated variance–covariance matrix from the SUR estimator.

([ ] out of 1p) Q2.f: Decision

At the 5% significance level, we reject the joint null hypothesis that all coefficients on `treat` are equal to zero across the system of equations. —

([ ] out of 3p) Q3: Implement Procedure 5 (Hotelling's T2 Test)

Script and Output

```r
opvs <- c("age","edu","nodegree","black","hisp",
          "married","u74","u75","re74","re75")

X1 <- df[df$treat == 1, opvs]
X0 <- df[df$treat == 0, opvs]

# Use DescTools::HotellingsT2Test()
if (!requireNamespace("DescTools", quietly = TRUE)) install.packages("DescTools")
library(DescTools)
```

```
## Warning: package 'DescTools' was built under R version 4.5.1
```

```
##
## Attaching package: 'DescTools'
## The following object is masked from 'package:car':
##
##     Recode
```

```
ht <- DescTools::HotellingsT2Test(X1, X0)
ht
```

```
##
##  Hotelling's two sample T2-test
##
## data:  X1 and X0
## T.2 = 2.005, df1 = 10, df2 = 434, p-value = 0.0314
## alternative hypothesis: true location difference is not equal to c(0,0,0,0,0,0,0,0,0,0)
```

```
# Extract T2 and p-value (in case you want to print neatly)
T2 <- as.numeric(ht$statistic)
p_T2 <- as.numeric(ht$p.value)

c(T2 = T2, p_value = p_T2)
```

```
##         T2    p_value
## 2.00498455 0.03140293
```

Commentary

We implement Procedure 5 using Hotelling's two-sample T^2 test via `DescTools::HotellingsT2Test()` to assess joint balance in the $J = 10$ observed predetermined variables (OPVs). The null hypothesis is

$$H_0: \ \mathbb{E}[X \mid D = 1] = \mathbb{E}[X \mid D = 0],$$

against the alternative that at least one component differs.

The test yields a Hotelling's statistic of $T^2 = 2.005$. Using the finite-sample F-approximation with degrees of freedom $(10, 434)$, the associated p-value is 0.0314. At the 5% significance level, we therefore reject the null hypothesis of joint balance in the OPVs.

This result is consistent with the joint SUR-based Wald test in Procedure~4 and with the treatment-assignment regression test in Procedure~6. Differences in numerical values across procedures reflect different finite-sample approximations (chi-square versus F), not differences in the underlying hypothesis being tested. —

([ ] out of 4p) Q4: Implement Procedure 6 (OPV do not predict treatment assignment)

Script and Output

```
opvs <- c("age","edu","nodegree","black","hisp",
          "married","u74","u75","re74","re75")

lm_fit <- lm(treat ~ age + edu + nodegree + black + hisp +
               married + u74 + u75 + re74 + re75,
             data = df)

summary(lm_fit)$r.squared
```

```
## [1] 0.04415781
```

```
summary(lm_fit)$fstatistic   # (value, numdf, dendf)
```

```
##      value      numdf      dendf
##   2.004985  10.000000 434.000000
```

```
R2 <- summary(lm_fit)$r.squared
n  <- nobs(lm_fit)
M  <- length(opvs)  # number of slope parameters (10)

F_manual <- (R2 / M) / ((1 - R2) / (n - (M + 1)))
p_manual <- 1 - pf(F_manual, df1 = M, df2 = n - (M + 1))

c(R2 = R2, n = n, M = M, F_manual = F_manual, p_manual = p_manual)
```

```
##           R2            n            M     F_manual     p_manual
##   0.04415781 445.00000000  10.00000000   2.00498455   0.03140293
```

Commentary

We implement Procedure 6 by regressing the treatment indicator on a constant and the $J = 10$ observed predetermined variables using a linear probability model. Let R^2 denote the coefficient of determination from this regression and let $M = 10$ be the number of slope parameters. Following the companion document, we compute the overall significance test statistic

$$F = \frac{R^2/M}{(1 - R^2)/(n - (M+1))},$$

and obtain the corresponding p-value from an $F_{M,\,n-(M+1)}$ distribution. Using $n = 445$, we obtain $F = 2.005$ with a p-value of 0.0314. We reject the null hypothesis. —

([ ] out of 5p) Q5: Implement Procedures 7 and 8 (control for FWER)

Script and Output

```
# Q5: Procedures 7 (Bonferroni) and 8 (Holm-Bonferroni)
# We use the one-at-a-time p-values for H0,j: pi_{1,j}=0 (treat coefficient = 0) for each OPV.

opvs <- c("age","edu","nodegree","black","hisp","married","u74","u75","re74","re75")

# If you already created ols_fits in Q2.b, you can reuse it.
# Otherwise:
ols_fits <- lapply(opvs, function(v) lm(as.formula(paste0(v, " ~ treat")), data = df))
names(ols_fits) <- opvs

# Extract unadjusted p-values for treat in each equation
p_raw <- sapply(ols_fits, function(m) summary(m)$coefficients["treat","Pr(>|t|)"])

# Procedure 7: Bonferroni adjusted p-values = p_adj_j = min(1, J * p_j)
# (implemented by p.adjust(method="bonferroni"))
p_bonf <- p.adjust(p_raw, method = "bonferroni")

# Procedure 8: Holm-Bonferroni adjusted p-values (implemented by p.adjust(method="holm"))
p_holm <- p.adjust(p_raw, method = "holm")

# Collect results
res_q5 <- data.frame(
  OPV = opvs,
  p_raw = as.numeric(p_raw),
  p_bonf = as.numeric(p_bonf),
  p_holm = as.numeric(p_holm),
  reject_bonf_5pct = (p_bonf <= 0.05),
  reject_holm_5pct = (p_holm <= 0.05)
)

# Show table sorted by raw p-values (helpful for interpretation)
res_q5[order(res_q5$p_raw), ]
```

```
##                 OPV       p_raw     p_bonf      p_holm reject_bonf_5pct
## nodegree nodegree 0.001398352 0.01398352 0.01398352             TRUE
## u75           u75 0.065468962 0.65468962 0.58922066            FALSE
## hisp         hisp 0.076473893 0.76473893 0.61179115            FALSE
## edu           edu 0.135411167 1.00000000 0.94787817            FALSE
## age           age 0.264764269 1.00000000 1.00000000            FALSE
## u74           u74 0.326208987 1.00000000 1.00000000            FALSE
## married   married 0.327408105 1.00000000 1.00000000            FALSE
## re75         re75 0.382253831 1.00000000 1.00000000            FALSE
## black       black 0.649493182 1.00000000 1.00000000            FALSE
## re74         re74 0.982318253 1.00000000 1.00000000            FALSE
##          reject_holm_5pct
## nodegree             TRUE
## u75                 FALSE
## hisp                FALSE
## edu                 FALSE
## age                 FALSE
## u74                 FALSE
## married             FALSE
## re75                FALSE
## black               FALSE
## re74                FALSE
```

Commentary

Q5 controls the family-wise error rate (FWER) when testing balance in the $J = 10$ OPVs using one-at-a-time tests of $H_{0,j} : \pi_{1,j} = 0$ for each OPV j. Following the programming guidance, we implement:

- **Procedure 7 (Bonferroni)**, which adjusts p-values as $p_j^{(B)} = \min\{1, Jp_j\}$; and
- **Procedure 8 (Holm–Bonferroni)**, a step-down procedure that is weakly less conservative than Bonferroni.

Both procedures are implemented using `stats::p.adjust()`. At the 5% significance level, only `nodegree` remains statistically significant after controlling the FWER under both Bonferroni and Holm–Bonferroni adjustments. All other OPVs have adjusted p-values well above 0.05.

Therefore, we reject the joint null hypothesis of perfect balance across the OPVs, but the evidence for imbalance is driven entirely by the `nodegree` variable. This result is consistent with the earlier joint balance tests (Procedures 4 and 5) and the treatment-assignment test (Procedure 6). —

([ ] out of 50p) PART II: Review of OLS for Causal Analysis

([ ] out of 3p) Q6: Two ways of obtaining the DM estimator

Script and Output

```
treated  <- read.csv("nswre74_treated.csv")
control  <- read.csv("nswre74_control.csv")
```

```r
treated$treat <- 1
control$treat <- 0

df <- rbind(treated, control)

df$re74 <- df$re74 / 1000
df$re75 <- df$re75 / 1000

## Keep / check core variables
vars <- c("re78", "treat", "re74", "re75")
stopifnot(all(vars %in% names(df)))

## Optional: quick checks
with(df, table(treat))
```

```
## treat
##   0   1
## 260 185
```

```r
summary(df[, vars])
```

```
##       re78           treat             re74              re75
##  Min.   :    0   Min.   :0.0000   Min.   : 0.0000   Min.   : 0.000
##  1st Qu.:    0   1st Qu.:0.0000   1st Qu.: 0.0000   1st Qu.: 0.000
##  Median : 3702   Median :0.0000   Median : 0.0000   Median : 0.000
##  Mean   : 5301   Mean   :0.4157   Mean   : 2.1023   Mean   : 1.377
##  3rd Qu.: 8125   3rd Qu.:1.0000   3rd Qu.: 0.8244   3rd Qu.: 1.221
##  Max.   :60308   Max.   :1.0000   Max.   :39.5707   Max.   :25.142
```

```r
df$re74 <- df$re74 / 1000
df$re75 <- df$re75 / 1000

fit1 <- lm(re78 ~ treat, data = df)
rho_ols <- coef(fit1)[["treat"]]

dm <- with(df, mean(re78[treat == 1]) - mean(re78[treat == 0]))

# optional: "manual OLS" slope via covariance/variance
rho_manual <- with(df, cov(treat, re78) / var(treat))

c(rho_ols = rho_ols, dm = dm, rho_manual = rho_manual)
```

```
##    rho_ols         dm rho_manual
##   1794.342   1794.342   1794.342
```

```r
all.equal(rho_ols, dm)
```

```
## [1] TRUE
```

```r
all.equal(rho_ols, rho_manual)
```

```
## [1] TRUE
```

Commentary

Estimate of spec 1 by OLS, regressing post-intervention earnings on the treatment indicator is found by:

$$\texttt{re78}_i = \alpha + \rho, \texttt{treat}_i + u_i.$$

Under the random assignment of the NSW experiment, the OLS estimator of ρ coincides with the difference in mean outcomes between treated and control groups. To verify this result, we compute:

1. $\hat{\rho}_{\text{OLS}}$: the OLS slope on $\texttt{treat}_i$;
2. $\widehat{DM} = \bar{\texttt{re78}} * 1 - \bar{\texttt{re78}} * 0$: the difference in sample average earnings between treated and control;
3. $\hat{\rho}_{\text{manual}}$: the OLS slope computed manually as $\widehat{\text{Cov}}(\texttt{treat}_i, \texttt{re78}_i)/\widehat{\text{Var}}(\texttt{treat}_i)$ (to be sure)

Results:

$$\hat{\rho}_{\text{OLS}} = 1794.342, \qquad \widehat{DM} = 1794.342, \qquad \hat{\rho}_{\text{manual}} = 1794.342.$$

As seen above, all quantities are numerically identical. Indeed, the OLS estimator of the treatment coefficient exactly equals the difference in average 1978 earnings between treated and control units. This was checked by the manual computation in $\hat{\rho}_{\text{manual}}$. The estimated ATE implies that being offered training increases earnings in 1978 by approximately $1,794 on average.

([ ] out of 4p) Q7: Fully-saturated specification in Nodegree

Script and Output

```r
df$degree <- 1 - df$nodegree

fit_fs <- lm(re78 ~ 0 + nodegree + degree + treat:nodegree + treat:degree, data = df)
coef(fit_fs)
```

```
##       nodegree         degree nodegree:treat   degree:treat 
##       4495.415       4854.493       1154.047       3192.025
```

```r
b <- coef(fit_fs)

rho1 <- b["nodegree:treat"]   # CATE for nodegree = 1
rho2 <- b["degree:treat"]     # CATE for nodegree = 0

s1 <- mean(df$nodegree == 1)  # share without degree
s0 <- mean(df$nodegree == 0)  # share with degree

ate_hat <- s1 * rho1 + s0 * rho2

c(
  CATE_nodegree1 = rho1,
  CATE_nodegree0 = rho2,
  ATE = ate_hat
)
```

```
## CATE_nodegree1.nodegree:treat   CATE_nodegree0.degree:treat 
##                      1154.047                      3192.025 
##             ATE.nodegree:treat 
##                      1598.281
```

([ ] out of 2p) Q8: Get $\widehat{ATE}$ in one step

Implementing the $M = 2$ version of the shortcut specification, we have:

Script and Output

```r
df$re74 <- df$re74 / 1000
df$re75 <- df$re75 / 1000

B2 <- as.integer(df$nodegree == 1)  # nodegree group
B1 <- 1 - B2                        # degree group

N  <- nrow(df)
N2 <- sum(B2)
N1 <- sum(B1)

Z_eps <- df$treat * B2 * (N2 / N)
Z_varpi1 <- df$treat * (B1 - B2 * (N1 / N2))

fit_short <- lm(df$re78 ~ 0 + B1 + B2 + Z_varpi1 + Z_eps)
coef(fit_short)[["Z_eps"]]
```

```
## [1] 2613.452
```

The coefficient on `Z_eps` is $\hat{\epsilon}$ which equals $\widehat{ATE}$ by construction.

([ ] out of 22p) Q9: Implications of estimating a not-fully saturated specification in an OPV that takes M distinct values

In the fully saturated approach, we have $\hat{\rho}_m = \bar{Y}_{1,m} - \bar{Y}_{0,m}$. Consider instead the shortcut regression

$$Y_i = \sum_{m=1}^{M} \theta_m 1[x_i = a_m] + \rho D_i + u_i.$$

Partialling out the indicators $1[x_i = a_m]$ yields residualized variables $\tilde{Y}_i = Y_i - \bar{Y}_{m(i)}, \tilde{D}_i = D_i - \bar{D}_{m(i)}$, hence

$$\hat{\rho} = \frac{\sum_i \tilde{D}_i \tilde{Y}_i}{\sum_i \tilde{D}_i^2}.$$

Grouping by m:

$$\hat{\rho} = \sum_{m=1}^{M} \hat{\omega}_m \big(\bar{Y}_{1,m} - \bar{Y}_{0,m}\big), \qquad \hat{\omega}_m = \frac{\hat{s}_m \widehat{Var}(D|x = a_m)}{\sum_{j=1}^{M} \hat{s}_j \widehat{Var}(D|x = a_j)},$$

where $\widehat{Var}(D|x = a_m) = \bar{D}_m(1 - \bar{D}_m)$.

To investigate when we have $\hat{\rho} = \widehat{ATE}$, we have that the fully saturated ATE estimator is $\widehat{ATE} = \sum_m \hat{s}_m \hat{\rho}_m$, while the shortcut is $\hat{\rho} = \sum_m \hat{\omega}_m \hat{\rho}_m)$. Thus we have $\hat{\rho} = \widehat{ATE}$ if either:

1. treated shares are identical across m (so $\widehat{Var}(D|m)$ is constant and $\hat{\omega}_m = \hat{s}_m$); or
2. $\hat{\rho}_m$ is identical across m, so any weighting yields the same average.

Proving the consistency of $\hat{\rho}$, the deck states that regressing Y_i on $(1, D_i, 1[x_i = a_1], \ldots, 1[x_i = a_M])$ yields an OLS coefficient on D_i that is consistent for ATE iff assignment probabilities are independent of x_i or treatment effects are homogeneous in x_i as in the notes. This matches the limit of the weighting representation: the shortcut converges to a variance-weighted average of CATEs, which equals the population-share-weighted ATE only under those conditions.

Showing even split and precision weighting, if $\hat{s}_m = 1/M$ for all m, then $\widehat{ATE} = \frac{1}{M}\sum_m \hat{\rho}_m$, namely equal weight across cells.

Pertaining to the shortcut, we have that $\hat{\omega}_m \propto Var(D|m) = \bar{D}_m(1 - \bar{D}_m)$, so it emphasizes cells with more within-cell treated/control balance. Under the usual within-cell homoskedasticity condition, slope variance scales like $\sigma_m^2/SST_{D,m}$, and for binary (D), $SST_{D,m} \propto n_m\bar{D}_m(1 - \bar{D}_m)$ (Companion-to-NSW-Application: derivation using `SST_D` and $n\bar{D}(1 - \bar{D})$). Hence larger $Var(D|m)$ corresponds to more precise within-cell treatment-effect estimation, so the shortcut places relatively more weight on the more precisely estimated CATEs in that sense.

Takeaway: Fully saturating in a discrete confounder x averages cell-by-cell treatment effects using the population (or sample) composition. The fixed-effects shortcut that adds x indicators but does not interact them with treatment instead averages the same cell-by-cell effects using weights that depend on how much within-cell variation there is in treatment assignment. This generally changes the estimand unless either assignment rates are constant across cells or treatment effects are constant across cells.

([ ] out of 14p) Q10: Implement the DM estimator with Regression Adjustment (various specifications)

([ ] out of 3p) Q10.a: Estimate Specifications 2, 3 and 4

([ ] out of 1p) Q10.a.i: Estimate Specification 2

Script and Output

```r
df$re74 <- df$re74 / 1000
df$re75 <- df$re75 / 1000

# Spec 2
fit2 <- lm(re78 ~ treat + nodegree + edu, data = df)

# Spec 3 (nodegree + edu + other 8 OPVs; typical NSW OPVs shown)
fit3 <- lm(re78 ~ treat + nodegree + edu + age + black + hisp + married + re74 + re75 + u74 + u75,
           data = df)

# Spec 4: add treat × (age - mean(age))
df$age_c <- df$age - mean(df$age)
fit4 <- lm(re78 ~ treat + nodegree + edu + age + black + hisp + married + re74 + re75 + u74 + u75 +
             treat:age_c,
           data = df)

# ATE estimate + t-test (via lm summary)
summ <- function(fit) summary(fit)$coef["treat", c("Estimate","Std. Error","t value","Pr(>|t|)")]
rbind(Spec2 = summ(fit2), Spec3 = summ(fit3), Spec4 = summ(fit4))
```

```
##       Estimate Std. Error  t value    Pr(>|t|)
## Spec2 1645.927   638.0410 2.579657 0.010212514
## Spec3 1670.709   641.1321 2.605873 0.009479869
## Spec4 1659.602   641.2881 2.587920 0.009981085
```

```r
summ(fit2)
```

```
##     Estimate   Std. Error      t value     Pr(>|t|)
## 1.645927e+03 6.380410e+02 2.579657e+00 1.021251e-02
```

Commentary

The table above reports estimates of the ATE of being offered training on 1978 earnings under three regression-adjustment specifications. Across Specifications 2–4, the estimated ATE is remarkably stable, ranging from approximately $1,646 to $1,671. In all cases, the null hypothesis that the ATE is zero is rejected at the 1% level using the standard OLS t-test.

Interpretation: Spec 2 adjusts only for educational attainment (`nodegree`, `edu`). The estimated ATE is positive and statistically significant at the 5% level. This indicates that, even with a minimal set of controls focused on education, the offer of training is associated with an increase in 1978 earnings of roughly $1,646 on average. Because education is a strong predictor of earnings, including it helps account for outcome variation without materially altering the experimental contrast between treated and control units.

([ ] out of 1p) Q10.a.ii: Estimate Specification 3

Script and Output

```
summ(fit3)
```

```
##     Estimate   Std. Error      t value     Pr(>|t|)
## 1.670709e+03 6.411321e+02 2.605873e+00 9.479869e-03
```

Commentary

Spec 3 adds the full set of observed pre-treatment variables to the regression. The estimated ATE remains very close to that from Spec 2 and is again statistically significant at the 1% level. The similarity of the point estimates across Specifications 2 and 3 suggests that additional OPVs beyond education do not meaningfully change the estimated treatment effect, consistent with the plausibility of random assignment. The standard error is also of similar magnitude, indicating modest efficiency gains from the expanded covariate set.

([ ] out of 1p) Q10.a.iii: Estimate Specification 4

Script and Output

```
summ(fit4)
```

```
##     Estimate   Std. Error      t value     Pr(>|t|)
## 1.659602e+03 6.412881e+02 2.587920e+00 9.981085e-03
```

Commentary

Spec 4 allows for treatment-effect heterogeneity by age through an interaction between treatment and centered age. The estimated average effect, evaluated at the sample mean age, is again very close to those obtained under Specifications 2 and 3 and remains statistically significant at conventional levels. This indicates that allowing for age-related heterogeneity does not materially alter the estimated average impact of the training offer, reinforcing the robustness of the ATE estimate across alternative regression-adjustment specifications.

Across all three specifications, the estimated ATE is stable—approximately $1,650–$1,670—and statistically significant. This consistency suggests that regression adjustment using balanced OPVs primarily serves to refine precision rather than to change the estimated treatment effect.

([ ] out of 5p) Q10.b: Reasons to include OPVs when they are balanced

Even if OPVs are balanced between treated and control groups, including them as regression covariates can be useful for at least two reasons:

1. If OPVs explain variation in `re78`, then controlling for them can reduce the residual variance and thereby reduce the standard error of $\hat{\rho}$, yielding higher precision and tighter inference for the ATE. This is relevant even under randomized assignment because balance affects bias, not necessarily variance.
2. Even in experiments, exact balance typically holds only in expectation. In a realized finite sample, adjusting for prognostic covariates can stabilize estimates and improve finite-sample efficiency.

Moving from Spec 2 to Spec 3 adds additional OPVs. If these added OPVs are predictive of `re78`, then we would expect the standard error of the `treat` coefficient to decrease or , while the point estimate may or may not change depending on finite-sample correlations.

([ ] out of 1p) Q10.c: Is it problematic to regression-adjust for OPVs that are lagged outcomes?

Using lagged outcomes (as is the case for `re74`, `re75`) as regression covariates is not problematic so long as they are measured prior to treatment. In that case they are OPVs (pre-treatment variables) and can help improve precision because they tend to be strong predictors of future earnings. The main concern would be conditioning on variables affected by treatment (post-treatment variables), but `re74` and `re75` are pre-intervention in the NSW setting.

([ ] out of 5p) Q10.d: Interactions of OPV with Treatment Indicator and Two Hypothesis Testing Problems

Script and Output

```r
library(car)

# (i) Test H0: ATE = 0 in Spec 4 (coefficient on treat)
car::linearHypothesis(fit4, "treat = 0")
```

```
##
## Linear hypothesis test:
## treat = 0
##
## Model 1: restricted model
## Model 2: re78 ~ treat + nodegree + edu + age + black + hisp + married +
##     re74 + re75 + u74 + u75 + treat:age_c
##
##   Res.Df        RSS Df Sum of Sq      F   Pr(>F)
## 1    433 1.8634e+10
## 2    432 1.8349e+10  1 284471777 6.6973 0.009981 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# (ii) Test H0: effect does not vary by age (interaction term equals 0)
car::linearHypothesis(fit4, "treat:age_c = 0")
```

```
##
## Linear hypothesis test:
## treat:age_c = 0
##
## Model 1: restricted model
## Model 2: re78 ~ treat + nodegree + edu + age + black + hisp + married +
##     re74 + re75 + u74 + u75 + treat:age_c
##
##   Res.Df        RSS Df Sum of Sq      F Pr(>F)
## 1    433 1.8389e+10
## 2    432 1.8349e+10  1  39465057 0.9291 0.3356
```

Commentary

Interactions between OPVs and the treatment indicator allow the conditional average treatment effect to vary with observed characteristics (treatment effect heterogeneity). In Spec 4, the interaction with centered age explicitly models heterogeneity by age.

In each hypothesis test, we have

1. (H_0:) ATE is zero corresponds to testing $\rho = 0$, so `treat = 0`.
2. (H_0:) no age heterogeneity corresponds to testing $\gamma = 0$, so `treat:age_c = 0`.

Each hypothesis is tested one-at-a-time using `car::linearHypothesis()` as instructed in the guidance.

([ ] out of 5p) Q11: Mechanisms for NSW intervention to impact post-intervention earnings

A plausible mechanism is that being offered the NSW program increases earnings in 1978 by increasing employment and job-specific human capital: the offer provides access to subsidized employment and training, which can raise labor market experience and skills, improving productivity and employability, and thus increasing post-intervention earnings measured in 1978.