Example 4a: General Linear Models with Multiple Fixed Effects of Multiple Predictors Simultaneously *(complete syntax, data, and output available for STATA, R, and SAS electronically)*

The data for this example were selected from the 2012 General Social Survey dataset (and were also used for Examples 1, 2, and 3). Building on these prior examples, this example will examine the unique effects of three-category working class, linear and quadratic slopes for years of age, and three piecewise slopes (i.e., linear splines) for years of education in predicting annual income. It will also demonstrate how the results from hierarchical (stepwise) regression can be obtained from a single model using multivariate Wald F-tests instead, as well as how to compute effect sizes per fixed slope (or linear combinations thereof).

<u>STATA</u> Syntax for Importing and Preparing Data for Analysis:

```
// Paste in the folder address where "GSS_Example.xlsx" is saved between " "
cd "\\Client\C:\Dropbox\24_PSQF6243\PSQF6243_Example4a"

// IMPORT GSS_Example.xlsx data from working directory using exact file name
// To change all variable names to lowercase, remove "case(preserve")
clear // Clear before means close any open data
import excel "GSS_Example.xlsx", case(preserve) firstrow clear

// Create and label predictor variables for model 1 (same as in Example 2)
// Linear education predictor centered so that 0 is meaningful
gen educ12=educ-12
label variable educ12 "educ12: Education (0=12 years)"
// Recode marry predictor so that 0 is meaningful
gen marry01=. // Create new empty variable, then recode
replace marry01=0 if marry==1
replace marry01=1 if marry==2
label variable marry01 "marry01: 0=unmarried, 1=married"

// Create and label predictor variables for model 2 (same as in Example 3)
// 2 Indicator-dummy-coded binary predictors for workclass
gen LvM=. // Make two new empty variables
gen LvU=.
replace LvM=0 if workclass==1 // Replace each for lower
replace LvU=0 if workclass==1
replace LvM=1 if workclass==2 // Replace each for middle
replace LvU=0 if workclass==2
replace LvM=0 if workclass==3 // Replace each for upper
replace LvU=1 if workclass==3
label variable LvM "LvM: Low=0 v Mid=1 Class"
label variable LvU "LvU: Low=0 v Upp=1 Class"

// Center age at 18 (minimum in sample)
gen age18 = age-18
label variable age18 "age18: Age (0=18 years)"

// 3 Piecewise slopes for education
gen lessHS=. // Make 3 new empty variables
gen gradHS=.
gen overHS=.
// Replace for educ less than 12
replace lessHS=educ-11 if educ <  12
replace gradHS=0 if educ <  12
replace overHS=0 if educ <  12
// Replace for educ greater or equal to 12
replace lessHS=0 if educ >= 12
replace gradHS=1 if educ >= 12
replace overHS=educ-12 if educ >= 12
// Label variables
label variable lessHS "lessHS: Slope for Years Ed Less Than High School"
label variable gradHS "gradHS: Acute Bump for Graduating High School"
label variable overHS "overHS: Slope for Years Ed After High School"
```

```
// Label outcome
label variable income "income: Annual Income in 1000s"
```

R Syntax for Importing and Preparing Data for Analysis
(after loading packages *readxl*, *supernova*, *multcomp*, *lmhelpers*, and *TeachingDemos*):

```
# Set working directory (to import and export files to)
# Paste in the folder address where "GSS_Example.xlsx" is saved in quotes
setwd("C:/Dropbox/24_PSQF6243/PSQF6243_Example4a")

# Import GSS_Example.xlsx data from working directory -- path = file name
Example4a = read_excel(path="GSS_Example.xlsx", sheet="GSS_Example")
# Convert to data frame to use for analysis
Example4a = as.data.frame(Example4a)

### Create and label predictor variables for model 1 (same as in Example 2)
# Linear predictor for education centered so that 0 is meaningful
Example4a$educ12=Example4a$educ-12
# educ12: Education (0=12 years)
# Recode marry predictor so that 0 is meaningful
Example4a$marry01=NA  # Create new empty variable, then recode
Example4a$marry01[which(Example4a$marry==1)]=0
Example4a$marry01[which(Example4a$marry==2)]=1
# marry01: 0=unmarried, 1=married

### Create and label predictor variables for model 2 (same as in Example 3)
# 2 Indicator-dummy-coded binary predictors for workclass
Example4a$LvM=NA; Example4a$LvU=NA  # Make 2 new empty variables
Example4a$LvM[which(Example4a$workclass==1)]=0  # Replace each for lower
Example4a$LvU[which(Example4a$workclass==1)]=0
Example4a$LvM[which(Example4a$workclass==2)]=1  # Replace each for middle
Example4a$LvU[which(Example4a$workclass==2)]=0
Example4a$LvM[which(Example4a$workclass==3)]=0  # Replace each for upper
Example4a$LvU[which(Example4a$workclass==3)]=1
# LvM: Low=0 vs Mid=1 Class
# LvU: Low=0 vs Upp=1 Class

# Center age at 18 (minimum in sample)
Example4a$age18=Example4a$age-18
# age18: Age (0=18 years)
# 3 Piecewise slopes for education

# Make 3 new empty variables
Example4a$lessHS=NA; Example4a$gradHS=NA; Example4a$overHS=NA
# Replace each for educ less than 12
Example4a$lessHS[which(Example4a$educ<12)]=Example4a$educ[which(Example4a$educ<12)]-11
Example4a$gradHS[which(Example4a$educ<12)]=0
Example4a$overHS[which(Example4a$educ<12)]=0
# Replace each for educ greater or equal to 12
Example4a$lessHS[which(Example4a$educ>=12)]=0
Example4a$gradHS[which(Example4a$educ>=12)]=1
Example4a$overHS[which(Example4a$educ>=12)]=Example4a$educ[which(Example4a$educ>=12)]-12
# lessHS: Slope for Years Ed Less Than High School
# gradHS: Acute Bump for Graduating High School
# overHS: Slope for Years Ed After High School
# Label outcome variable
# income: Annual Income in 1000s
```

Note: I also wrote five custom functions to automate calculations of effect sizes from lm or glht output—please see code online for these (as used in this example).

Model 1: Linear Education and Binary Marital Status Predicting Income

Below is a summary of the results from estimating separate models per predictor (as in Example 2). Because there was only one slope in each model, the model R^2 = semi-partial R^2 = partial R^2 (see excel sheet online).

Separate Models	DF num	SS effect	SS residual	SS total	semi-partial R2 = SS effect / SS total	partial R2 = SS effect / (SS effect + SS residual)
Empty Model	0	0	139423	139423	0.0000	0.0000
Linear Education	1	20635	118788	139423	0.1480	0.1480
Binary Marital Status	1	7060	132363	139423	0.0506	0.0506
Sum of Separate Models	2	27695			0.1986	0.1986

Combined: $\boldsymbol{Income_i = \beta_0 + \beta_1(educ_i - 12) + \beta_2(marry01_i) + e_i}$

```
display "STATA Model 1: Linear Education and Binary Marital Status"
regress income c.educ12 c.marry01, level(95)

      Source |       SS           df       MS      Number of obs   =       734
-------------+----------------------------------   F(2, 731)       =     85.89
       Model |  26530.4118         2  13265.2059   Prob > F        =    0.0000
    Residual |   112892.82       731  154.436142   R-squared       =    0.1903
-------------+----------------------------------   Adj R-squared   =    0.1881
       Total |  139423.232       733  190.209048   Root MSE        =    12.427
```

Mean Square Residual / Error, the residual variance, is 154.44 after including 1 slope each for the 2 predictor constructs (which accounted for 19.03% of the variance in income as the model R^2 = .1903). The *F*-test says this R^2 is significantly > 0, $F(2, 731) = 85.89$, MSE = 154.44, $p < .001$.

```
------------------------------------------------------------------------------
      income |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      educ12 |    1.77385    .157981    11.23   0.000     1.463699       2.084  beta1
     marry01 |   5.694607   .9216806     6.18   0.000      3.88515    7.504064  beta2
       _cons |   11.47412   .6775219    16.94   0.000       10.144    12.80425  beta0
------------------------------------------------------------------------------

matrix Model1 = r(table)   // Save results for computing effect sizes
matrix list Model1         // Show saved results    d = 2t/SQRT(DFden)

Model1[9,3]
           educ12     marry01       _cons
     b  1.7738496   5.6946068   11.474125
    se  .15798099   .92168064   .67752192
     t  11.228247   6.1785032   16.935429
pvalue  4.230e-27   1.074e-09   1.597e-54
    ll   1.463699     3.88515   10.144004
    ul  2.0840002   7.5040636   12.804246
    df        731         731         731
  crit  1.9632145   1.9632145   1.9632145
 eform          0           0           0

display "Partial Cohen's D for marry01 = " 2*Model1[3,2]/sqrt(Model1[7,2])
Partial Cohen's D for marry01 = .45704039
```

A separate STATA command, pcorr, provides partial and semi-partial effect sizes for fixed slopes:

```
display "STATA Partial and Semi-Partial R and R2 Effect Sizes for Fixed Slopes"
pcorr income c.educ12 c.marry01
```

```
Partial and semipartial correlations of income with
                   Partial   Semipartial      Partial   Semipartial   Significance
   Variable |       Corr.         Corr.      Corr.^2       Corr.^2          Value
------------+-----------------------------------------------------------------
     educ12 |      0.3835        0.3737       0.1471        0.1396         0.0000
    marry01 |      0.2228        0.2056       0.0496        0.0423         0.0000
```

The **two most useful effect sizes** can be interpreted as follows:
Partial r = effect size in correlation metric for unique effect controlling for other predictors
Semi-partial R^2 = proportion of explained variance (amount of model R^2) attributable to that fixed slope

---

The R code and output below uses three custom functions to organize output and compute effect sizes:

```
print("R Model 1: Linear Education and Binary Marital Status")
Model1 = lm(data=Example4a, formula=income~1+educ12+marry01)
supernova(Model1)             # supernova prints sums of squares and residual variance

Analysis of Variance Table (Type III SS)
 Model: income ~ 1 + educ12 + marry01
                                       SS  df        MS       F   PRE     p
 ------- --------------- | ---------- --- --------- ------- ----- -----
   Model (error reduced) |  26530.412   2 13265.206  85.894 .1903 .0000
  educ12                 |  19470.310   1 19470.310 126.074 .1471 .0000
 marry01                 |   5895.430   1  5895.430  38.174 .0496 .0000
   Error (from model)    | 112892.820 731   154.436
 ------- --------------- | ---------- --- --------- ------- ----- -----
   Total (empty model)   | 139423.232 733   190.209
```

PRE gives partial R^2 (so model version is full R^2)

Mean Square Residual / Error, the residual variance, is 154.44 after including 1 slope each for the 2 predictor constructs (which accounted for 19.03% of the variance in income as the model R^2 = .1903). The *F*-test says this R^2 is significantly > 0, $F(2, 731) = 85.89$, MSE = 154.44, $p < .001$.

```
SummaryCI(Model1, level=.95)  # custom function to add CIs to fixed effects table

             Estimate  StdErr t-value    p-value LowerCI UpperCI
(Intercept)   11.4741 0.67752 16.9354 1.5965e-54 10.1440 12.8042   beta0
educ12         1.7738 0.15798 11.2282 4.2301e-27  1.4637  2.0840   beta1
marry01        5.6946 0.92168  6.1785 1.0743e-09  3.8851  7.5041   beta2

FixedEffectSizes(Model1)      # custom function to add effect sizes for fixed slopes

             Estimate    p-value Partial-D Partial-R SemiPartial-R Partial-R2 SemiPartial-R2
(Intercept)   11.4741 1.5965e-54   1.25276   0.53084       0.56364    0.28179       0.317692
educ12         1.7738 4.2301e-27   0.83058   0.38353       0.37370    0.14710       0.139649
marry01        5.6946 1.0743e-09   0.45704   0.22278       0.20563    0.04963       0.042284
```

Below is a summary of the results for the overall model and the contribution of each predictor— As shown in the last row above, the sum across the two constructs of the Effects Sums of Squares (SS effect) differs from the Model SS (as SS effect for full model)—this is because the Model SS takes into account the predictive contribution of the shared variance among the sets of predictors (i.e., none of them "gets credit" for what they have in common that predicts income, but the model R^2 does reflect that common contribution).

Combined Model	**DF num**	**SS effect**	**SS residual**	**SS total**	**semi-partial R2 = SS effect / SS total**	**partial R2 = SS effect / (SS effect + SS residual)**
Full Model	2	26530	112893	139423	0.1903	0.1903
Linear Education	1	19470	112893	139423	0.1396	0.1471
Binary Marital Status	1	5895	112893	139423	0.0423	0.0496
Sum of Predictors		25366			0.1819	0.1967

Model 2: Three-Category Workclass, Linear and Quadratic Age Slopes, and Three Piecewise Linear Education Slopes Predicting Income

Below is a summary of the results from estimating separate models per conceptual predictor (as demonstrated in Example 3, as well as in the syntax and output online). Because there was only one conceptual predictor in each model, the model R^2 = semi-partial R^2 = partial R^2 (see excel sheet online).

Separate Models	DF num	SS effect	SS residual	SS total	semi-partial R2 = SS effect / SS total	partial R2 = SS effect / (SS effect + SS residual)
3-Category Workclass	2	14414	125009	139423	0.1034	0.1034
Linear + Quadratic Age	2	15885	123538	139423	0.1139	0.1139
3 Piecewise Slopes Education	3	22907	116517	139423	0.1643	0.1643
Sum of Separate Models	2	53206			0.3816	0.3816

Combined: $$Income_i = \beta_0 + \beta_1(LvsM_i) + \beta_2(LvsU_i) + \beta_3(Age_i - 18) + \beta_4(Age_i - 18)^2 + \beta_5(LessHS_i) + \beta_6(GradHS_i) + \beta_7(OverHS_i) + e_i$$

In addition to the overall F-test of the model R^2, the purpose of estimating a single model with the seven slopes from all three predictive constructs combined (workclass, age, and education) is to determine to what extent their bivariate effects (when each construct was in a separate model predicting income, as was the case in Example 3) differ from their unique effects (when all constructs are combined in the same model, below). The solution for the fixed effects will provide tests for the significance of each slope (against a null hypothesis of a 0 slope in the population), and we will also ask for joint F-tests (and their effect sizes) that combine the multiple slopes needed to capture the full effect of each construct. Effect sizes per slope are also reported below.

```
display "STATA Model 2: Three-Category Workclass, Quadratic Age, and Piecewise Education"
regress income c.LvM c.LvU c.age18 c.age18#c.age18 c.lessHS c.gradHS c.overHS, level(95)
```

```
      Source |       SS           df       MS      Number of obs   =       734
-------------+----------------------------------   F(7, 726)       =     42.09
       Model |  40246.4243         7  5749.48919   Prob > F        =    0.0000
    Residual |  99176.8076       726  136.607173   R-squared       =    0.2887
-------------+----------------------------------   Adj R-squared   =    0.2818
       Total |  139423.232       733  190.209048   Root MSE        =    11.688
```

Mean Square Residual / Error, the residual variance, is 136.61 after including 7 slopes for the 3 predictor constructs (which accounted for 28.87% of the variance in income as the model R^2 = .2887). The F-test says this R^2 is significantly > 0, $F(7, 726) = 42.04$, MSE = 136.61, $p < .001$.

```
------------------------------------------------------------------------------
          income |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----------------+------------------------------------------------------------
             LvM |   6.060105   .9470067     6.40   0.000     4.200907    7.919304   beta1
             LvU |   7.208538   2.697879     2.67   0.008     1.911961    12.50511   beta2
           age18 |    1.06998   .1230046     8.70   0.000     .8284928    1.311467   beta3
c.age18#c.age18 |  -.0175062   .0022749    -7.70   0.000    -.0219724    -.01304   beta4
          lessHS |   .2589179   .5612016     0.46   0.645    -.8428539    1.36069   beta5
          gradHS |   3.157139   1.757267     1.80   0.073    -.2927916    6.60707   beta6
          overHS |   1.528179   .2080423     7.35   0.000     1.119743    1.936616   beta7
           _cons |  -3.686546   2.004615    -1.84   0.066    -7.622081    .2489889   beta0
------------------------------------------------------------------------------
```

```
estimates store Model1        // Save all model results
ereturn list                  // See what has been stored automatically
```

```
scalars:
               e(rank) =  8
               e(ll_0) =  -2967.061946932051
                 e(ll) =  -2842.058095807061
```

```
              e(r2_a) =  .2818050735748676
               e(rss) =  99176.80756352992
               e(mss) =  40246.42433343873
              e(rmse) =  11.68790712456771
                e(r2) =  .2886636881519153
                 e(F) =  42.08775473663854
              e(df_r) =  726
              e(df_m) =  7
                 e(N) =  734

global SSresidual = e(rss)  // Save full model SS residual for effect sizes below
global SSfull = e(mss)      // Save full model SS model for effect sizes below

display "STATA Semi-Partial and Partial R and R2 Effect Sizes for Fixed Slopes"
pcorr income c.LvM c.LvU c.age18 c.age18#c.age18 c.lessHS c.gradHS c.overHS

Partial and semipartial correlations of income with

               Partial   Semipartial      Partial   Semipartial   Significance
   Variable |    Corr.         Corr.      Corr.^2       Corr.^2          Value
------------+-----------------------------------------------------------------
        LvM |   0.2311        0.2003       0.0534        0.0401         0.0000
        LvU |   0.0987        0.0836       0.0097        0.0070         0.0077
      age18 |   0.3072        0.2723       0.0944        0.0741         0.0000
c.age18#~18 |  -0.2746       -0.2409       0.0754        0.0580         0.0000
     lessHS |   0.0171        0.0144       0.0003        0.0002         0.6447
     gradHS |   0.0665        0.0562       0.0044        0.0032         0.0728
     overHS |   0.2630        0.2299       0.0692        0.0529         0.0000

display "STATA Get Model-Implied Difference for Middle vs Upper and Effect Sizes"
estimates restore Model1    // Restore model results for post-estimations below
lincom c.LvM*-1 + c.LvU*1   // Code below should work after any lincom
   display "Partial D= "      (2*(r(estimate)/r(se)))/sqrt(r(df))
   display "Partial R= "      (r(estimate)/r(se))/sqrt((r(estimate)/r(se))^2+r(df))
   display "Semi-Partial R= " (r(estimate)/r(se))*sqrt((1-e(r2))/r(df))

------------------------------------------------------------------------------
      income |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |   1.148432    2.70813     0.42   0.672    -4.168269    6.465134  beta2 – beta1
------------------------------------------------------------------------------
Partial D= .03147731
Partial R= .01573671
Semi-Partial R= .0132741
```

Because the workclass predictors are related (each shares a reference group with another), the total of the sr^2 values for these three differences they imply (two of which are given by pcorr; the other was requested as a linear combination) is greater than it should be. The linear age slope's sr^2 effect size is valid but conditional on age 18. The per-slope sr^2 effect sizes for the remainder of the slopes are ok, but these cannot be added together directly to represent the contribution per conceptual predictor. Instead, we need to obtain an F-test and effect size that combines the fixed slopes for the same construct… that's what the STATA code below gives us!

```
display "STATA F-test for Overall Effect of Workclass"
test (c.LvM=0)(c.LvU=0)
       F(  2,   726) =   21.82
            Prob > F =    0.0000

display "STATA F-test for Overall Effect of Age"
test (c.age18=0)(c.age18#c.age18=0)
       F(  2,   726) =   41.08
            Prob > F =    0.0000

display "STATA F-test for Overall Effect of Education"
test (c.lessHS=0)(c.gradHS=0)(c.overHS=0)
```

```
       F(  3,   726) =   27.46
            Prob > F =    0.0000

display "STATA Reduced Model to Get Model SS from Omitting Workclass"
quietly regress income c.age18 c.age18#c.age18 c.lessHS c.gradHS c.overHS, level(95)
global  SSeffect = $SSfull - e(mss)
display "Partial R2 = "      $SSeffect/($SSeffect+$SSresidual)
display "Semi-Partial R2 = " $SSeffect/($SSfull+$SSresidual)
Partial R2 = .0567038
Semi-Partial R2 = .04276013

display "STATA Reduced Model to Get Model SS from Omitting Age"
quietly regress income c.LvM c.LvU c.lessHS c.gradHS c.overHS, level(95)
global  SSeffect = $SSfull - e(mss)
display "Partial R2 = "      $SSeffect/($SSeffect+$SSresidual)
display "Semi-Partial R2 = " $SSeffect/($SSfull+$SSresidual)
Partial R2 = .10166273
Semi-Partial R2 = .08050027

display "STATA Reduced Model to Get Model SS from Omitting Education"
quietly regress income c.LvM c.LvU c.age18 c.age18#c.age18, level(95)
global  SSeffect = $SSfull - e(mss)
display "Partial R2 = "      $SSeffect/($SSeffect+$SSresidual)
display "Semi-Partial R2 = " $SSeffect/($SSfull+$SSresidual)
Partial R2 = .10189198
Semi-Partial R2 = .08070239
```

Below is a summary of the results for the overall model and the contribution of each predictor:

Combined Model	DF num	SS effect	SS residual	SS total	semi-partial R2 = SS effect / SS total	partial R2 = SS effect / (SS effect + SS residual)
Full Model	7	40246	99177	139423	0.2887	0.2887
2 Group Differences Workclass	2	5962	99177	139423	0.0428	0.0567
Linear + Quadratic Age	2	11224	99177	139423	0.0805	0.1017
3 Piecewise Slopes Education	3	11252	99177	139423	0.0807	0.1019
Sum of Workclass, Age, Education	7	28437			0.2040	0.2603

As shown in the last row above, the sum across the three constructs of the Effects Sums of Squares (SS effect) differs from the Model SS (SS effect from full model)—this is because the Model SS takes into account the predictive contribution of the shared variance among the sets of predictors (i.e., none of them "gets credit" for what they have in common that predicts income, but the model R^2 does reflect that common contribution).

```
print("R Model 2: Three-Category Workclass, Quadratic Age, and Piecewise Education")
Model2 = lm(data=Example4a, formula=income~1+LvM+LvU+age18+I(age18^2)+lessHS+gradHS+overHS)
supernova(Model2)    # supernova prints sums of squares and residual variance

Analysis of Variance Table (Type III SS)
 Model: income ~ 1 + LvM + LvU + age18 + I(age18^2) + lessHS + gradHS + overHS

                                    SS  df         MS      F   PRE     p
 ---------- --------------- | ---------- --- ---------- ------ ----- -----
      Model (error reduced) |  40246.424   7   5749.489 42.088 .2887 .0000
        LvM                 |   5594.069   1   5594.069 40.950 .0534 .0000
        LvU                 |    975.265   1    975.265  7.139 .0097 .0077
      age18                 |  10336.707   1  10336.707 75.667 .0944 .0000
 I(age18^2)                 |   8089.564   1   8089.564 59.218 .0754 .0000
     lessHS                 |     29.078   1     29.078  0.213 .0003 .6447
     gradHS                 |    440.946   1    440.946  3.228 .0044 .0728
     overHS                 |   7370.869   1   7370.869 53.957 .0692 .0000
      Error (from model)    |  99176.808 726    136.607
 ---------- --------------- | ---------- --- ---------- ------ ----- -----
      Total (empty model)   | 139423.232 733    190.209
```

Mean Square Residual / Error, the residual variance, is 136.61 after including 7 slopes for the 3 predictor constructs (which accounted for 28.87% of the variance in income as the model $R^2 = .2887$). The *F*-test says this R^2 is significantly > 0, $F(7, 726) = 42.04$, MSE = 136.61, $p < .001$.

```
SummaryCI(Model2, level=.95)  # custom function to add CIs to fixed effects table

               Estimate    StdErr  t-value    p-value   LowerCI  UpperCI
(Intercept) -3.686546 2.0046155 -1.83903 6.6319e-02 -7.622081  0.24899  beta0
LvM          6.060105 0.9470067  6.39922 2.7998e-10  4.200907  7.91930  beta1
LvU          7.208538 2.6978794  2.67193 7.7108e-03  1.911961 12.50511  beta2
age18        1.069980 0.1230046  8.69870 2.2305e-17  0.828493  1.31147  beta3
I(age18^2)  -0.017506 0.0022749 -7.69530 4.6148e-14 -0.021972 -0.01304  beta4
lessHS       0.258918 0.5612016  0.46136 6.4468e-01 -0.842854  1.36069  beta5
gradHS       3.157139 1.7572666  1.79662 7.2812e-02 -0.292792  6.60707  beta6
overHS       1.528179 0.2080423  7.34552 5.5215e-13  1.119743  1.93662  beta7


FixedEffectSizes(Model2)      # custom function to add effect sizes for fixed slopes

              Estimate    p-value Partial-D Partial-R SemiPartial-R Partial-R2 SemiPartial-R2
(Intercept) -3.686546 6.6319e-02 -0.136506 -0.068094     -0.057565  0.0046368     0.00331372
LvM          6.060105 2.7998e-10  0.474995  0.231070      0.200307  0.0533934     0.04012293
LvU          7.208538 7.7108e-03  0.198329  0.098681      0.083636  0.0097378     0.00699500
age18        1.069980 2.2305e-17  0.645678  0.307225      0.272285  0.0943875     0.07413906
I(age18^2)  -0.017506 4.6148e-14 -0.571199 -0.274619     -0.240877  0.0754157     0.05802164
lessHS       0.258918 6.4468e-01  0.034246  0.017120      0.014442  0.0002931     0.00020856
gradHS       3.157139 7.2812e-02  0.133358  0.066531      0.056237  0.0044264     0.00316265
overHS       1.528179 5.5215e-13  0.545236  0.263019      0.229928  0.0691791     0.05286686

print("R Ask for missing model-implied group difference")
glhtModel2 = glht(model=Model2, linfct=rbind("Mid vs Upp Diff" = c(0,-1,1,0,0,0,0,0)))
glhtSummaryCI(glhtModel2, level=.95)  # custom function to add CIs to glht output table

                 Estimate StdErr p-value LowerCI UpperCI
Mid vs Upp Diff    1.1484 2.7081 0.67164 -4.1683  6.4651  beta2 – beta1

glhtEffectSizes(glhtObject=glhtModel2, modelObject=Model2,
                level=.95) # custom function to compute glht effect sizes

                Estimate p-value Partial-D Partial-R SemiPartial-R Partial-R2 SemiPartial-R2
Mid vs Upp Diff   1.1484 0.67164  0.031477  0.015737      0.013274 0.00024764      0.0001762
```

Because the workclass predictors are related (each shares a reference group with another), the total of the sr^2 values for these three differences they imply (two of which are given by pcorr; the other was requested as a linear combination) is greater than it should be. The linear age slope's sr^2 effect size is valid but conditional on age 18. The per-slope sr^2 effect sizes for the remainder of the slopes are ok, but these cannot be added together directly to represent the contribution per conceptual predictor. Instead, we need to obtain an *F*-test and effect size that combines the slopes for the same construct… that's what the R code below gives us!

```
# Fit model without fixed slopes of interest (LvM and LvU for workclass here)
Model2NoClass = lm(data=Example4a, formula=income~1+age18+agesq+lessHS+gradHS+overHS)
# Get F-test and effect sizes for fixed slopes of interest using custom function
R2changeF(ReducedModel=Model2NoClass, FullModel=Model2, PredName="Workclass")

F-Test and R2 Change for Workclass Slopes
  R2-total R2-change DF-num DF-den F-value    p-value Partial-R2 SemiPartial-R2
2  0.28866   0.04276      2    726  21.821 6.2698e-10   0.056704        0.04276

# Repeat for age slopes
Model2NoAge = lm(data=Example4a, formula=income~1+LvM+LvU+lessHS+gradHS+overHS)
R2changeF(ReducedModel=Model2NoAge, FullModel=Model2, PredName="Age")
```

```
F-Test and R2 Change for Age Slopes
  R2-total R2-change DF-num DF-den F-value    p-value Partial-R2 SemiPartial-R2
2  0.28866    0.0805      2    726   41.08 1.2546e-17    0.10166         0.0805
```

```
# Repeat for education slopes
Model2NoEduc = lm(data=Example4a, formula=income~1+LvM+LvU+age18+agesq)
R2changeF(ReducedModel=Model2NoEduc, FullModel=Model2, PredName="Education")
```

```
F-Test and R2 Change for Education Slopes
  R2-total R2-change DF-num DF-den F-value  p-value Partial-R2 SemiPartial-R2
2  0.28866  0.080702      3    726  27.455 7.95e-17    0.10189       0.080702
```

Example Results Section for Model 2 (would continue from separate results described in Example 3):

Table 1

Final Model Results

Fixed Effect	Est	SE	p <	Cohen's *d*	Partial *r*
Intercept	−3.687	2.005	.001		
Lower vs Middle Class	6.060	0.947	.001	0.475	.231
Lower vs Upper Class	7.209	2.698	.008	0.198	.099
(Middle vs Upper Class)	1.148	2.708	.672	0.031	.016
Linear Age Slope	1.070	0.123	.001	0.646	.307
Quadratic Age Slope	−0.018	0.002	.001	−0.572	−.275
Education 2 to 11 years	0.259	0.561	.645	0.034	.017
Education: 11 to 12 years	3.157	1.757	.073	0.134	.067
Education: 12 to 20 years	1.528	0.208	.001	0.546	.263

Note: Cohen's *d* and partial *r* effect sizes were computed as: $d = \frac{2t}{\sqrt{DF_{den}}}$; $r = \frac{t}{\sqrt{t^2+DF_{den}}}$. Model-implied effects are given in parentheses, computed as linear combinations of the fixed effects.

After examining the bivariate contributions of three-category self-reported working class membership, linear and quadratic years of age, and piecewise slopes for years of education in separate models, we then estimated a combined model to examine their unique contributions after controlling for each other construct. Model 2 (including all seven fixed slopes) captured a significant amount of variance in annual income, $F(7, 726) = 42.09$, MSE = 136.61, $p < .001$, $R^2 = .289$. Parameter estimates and effect sizes are given in Table 1. Semipartial eta-squared (η^2) effect sizes and corresponding multivariate Wald *F*-tests were obtained to evaluate the amount of total variance captured by distinct sets of predictor slopes.

The omnibus unique effect of three-category self-reported working class membership remained significant, $F(2, 726) = 21.83$, MSE = 136.61, $p < .0001$, semipartial $\eta^2 = .043$. As shown in Table 1, relative to lower-class respondents (the reference group), after controlling for years of age and years of education, annual income was still significantly higher for both middle-class and upper-class respondents (by 6.060 and 7.209 thousand dollars, respectively). Middle-class and upper-class respondents still did not differ significantly in predicted annual income.

The omnibus unique effect of quadratic years of age (centered at 18) also remained significant, $F(2, 726) = 41.08$, MSE = 136.61, $p < .0001$, semipartial $\eta^2 = .081$. As shown in Table 1, after controlling for self-reported working class and years of education, annual income was expected to be significantly higher by 1.070 thousand dollars per year of age at age 18; this instantaneous linear age slope was predicted to become significantly less positive per year of age by twice the quadratic coefficient of −0.018. As given by the quantity (−1*linear slope) / (2*quadratic slope) + 18, the age of maximum predicted personal income was 48.56 (i.e., the age at which the linear age slope = 0).

The omnibus unique effect of piecewise years of education (centered at 11) also remained significant, $F(3, 726) = 27.46$, MSE = 136.61, $p < .0001$, semipartial $\eta^2 = .081$. As shown in Table 1, after controlling for self-reported working class and years of age, annual income was expected to be nonsignificantly higher by 0.259 thousand dollars per year of education from 2 to 11 years, to be nonsignificantly higher by 3.157 thousand dollars for those achieving a high school degree, and to be significantly higher by 1.528 thousand dollars per year of additional education past 12 years. Notably, the effect of a high school degree (the difference between 11 and 12 years of education) was no longer significant after controlling for age and self-reported working class membership.

[The rest of the text would need to emphasize why it matters based on your research questions that the predictors had significant unique effects. This is the part that must be customized per research study!]