SEM Preliminary

Ryan Larson - UMN

01/07/2020

```
ols <- lm(total_ff_log~black+hispanic+asian+nativeam+other.race+
          male+log(age)+priors+trial_flag+felony+gross.mis+mis+drug+alcohol.dui+hunt.fish+othe
          min_perc+repubvote+lecperc, data = monsanc.short)
summary(ols)
```

```
##
## Call:
## lm(formula = total_ff_log ~ black + hispanic + asian + nativeam +
##     other.race + male + log(age) + priors + trial_flag + felony +
##     gross.mis + mis + drug + alcohol.dui + hunt.fish + other.offense +
##     min_perc + repubvote + lecperc, data = monsanc.short)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.9694 -0.3673  0.1114  0.5529  3.2964
##
## Coefficients:
##                   Estimate Std. Error  t value Pr(>|t|)
## (Intercept)      5.264e+00  1.166e-02  451.572  < 2e-16 ***
## black           -1.141e-01  1.617e-03  -70.587  < 2e-16 ***
## hispanic         1.896e-02  2.541e-03    7.461 8.58e-14 ***
## asian            1.202e-02  3.222e-03    3.731 0.000191 ***
## nativeam        -6.360e-02  3.709e-03  -17.147  < 2e-16 ***
## other.race      -2.415e-02  4.592e-03   -5.259 1.45e-07 ***
## male             2.130e-02  1.374e-03   15.505  < 2e-16 ***
## log(age)        -1.693e-01  1.881e-03  -89.998  < 2e-16 ***
## priors          -7.877e-03  7.547e-05 -104.370  < 2e-16 ***
## trial_flagTRUE  -2.080e-01  7.650e-02   -2.719 0.006547 **
## felony          -9.394e-02  3.824e-03  -24.566  < 2e-16 ***
## gross.mis        5.497e-01  3.641e-03  150.977  < 2e-16 ***
## mis              5.083e-01  1.484e-03  342.461  < 2e-16 ***
## drug             4.025e-01  5.355e-03   75.165  < 2e-16 ***
## alcohol.dui      5.442e-01  4.807e-03  113.216  < 2e-16 ***
## hunt.fish       -1.902e-02  1.071e-02   -1.777 0.075647 .
## other.offense    2.619e-01  4.270e-03   61.325  < 2e-16 ***
## min_perc        -2.405e-02  1.757e-04 -136.888  < 2e-16 ***
## repubvote        1.544e-03  1.223e-04   12.627  < 2e-16 ***
## lecperc          4.636e-03  2.275e-04   20.382  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 1.071 on 2870581 degrees of freedom
##    (2691382 observations deleted due to missingness)
## Multiple R-squared:  0.1182, Adjusted R-squared:  0.1182
## F-statistic: 2.025e+04 on 19 and 2870581 DF,  p-value: < 2.2e-16
```

```r
#Multivariate Regression - when all predictors are identical across models
  #equivalent to separate OLS, but more efficient and takes into account error covariance

mv <- lm(cbind(total_ff_log, conf_days)~black+hispanic+asian+nativeam+other.race+
          male+log(age)+priors+trial_flag+felony+gross.mis+mis+drug+alcohol.dui+hunt.fish+othe
          min_perc+repubvote+lecperc, data = monsanc.short)
summary(mv)
```

```
## Response total_ff_log :
##
## Call:
## lm(formula = total_ff_log ~ black + hispanic + asian + nativeam +
##     other.race + male + log(age) + priors + trial_flag + felony +
##     gross.mis + mis + drug + alcohol.dui + hunt.fish + other.offense +
##     min_perc + repubvote + lecperc, data = monsanc.short)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.9694 -0.3673  0.1114  0.5529  3.2964
##
## Coefficients:
##                  Estimate Std. Error  t value Pr(>|t|)
## (Intercept)     5.264e+00  1.166e-02  451.572  < 2e-16 ***
## black          -1.141e-01  1.617e-03  -70.587  < 2e-16 ***
## hispanic        1.896e-02  2.541e-03    7.461 8.58e-14 ***
## asian           1.202e-02  3.222e-03    3.731 0.000191 ***
## nativeam       -6.360e-02  3.709e-03  -17.147  < 2e-16 ***
## other.race     -2.415e-02  4.592e-03   -5.259 1.45e-07 ***
## male            2.130e-02  1.374e-03   15.505  < 2e-16 ***
## log(age)       -1.693e-01  1.881e-03  -89.998  < 2e-16 ***
## priors         -7.877e-03  7.547e-05 -104.370  < 2e-16 ***
## trial_flagTRUE -2.080e-01  7.650e-02   -2.719 0.006547 **
## felony         -9.394e-02  3.824e-03  -24.566  < 2e-16 ***
## gross.mis       5.497e-01  3.641e-03  150.977  < 2e-16 ***
## mis             5.083e-01  1.484e-03  342.461  < 2e-16 ***
## drug            4.025e-01  5.355e-03   75.165  < 2e-16 ***
## alcohol.dui     5.442e-01  4.807e-03  113.216  < 2e-16 ***
## hunt.fish      -1.902e-02  1.071e-02   -1.777 0.075647 .
## other.offense   2.619e-01  4.270e-03   61.325  < 2e-16 ***
## min_perc       -2.405e-02  1.757e-04 -136.888  < 2e-16 ***
## repubvote       1.544e-03  1.223e-04   12.627  < 2e-16 ***
## lecperc         4.636e-03  2.275e-04   20.382  < 2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.071 on 2870581 degrees of freedom
##   (2691382 observations deleted due to missingness)
## Multiple R-squared:  0.1182, Adjusted R-squared:  0.1182
## F-statistic: 2.025e+04 on 19 and 2870581 DF,  p-value: < 2.2e-16
##
##
## Response conf_days :
##
## Call:
## lm(formula = conf_days ~ black + hispanic + asian + nativeam +
##     other.race + male + log(age) + priors + trial_flag + felony +
##     gross.mis + mis + drug + alcohol.dui + hunt.fish + other.offense +
##     min_perc + repubvote + lecperc, data = monsanc.short)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -1511     -8     -1      6  43562
##
## Coefficients:
##                 Estimate Std. Error  t value Pr(>|t|)
## (Intercept)     83.23286    2.38166   34.947   <2e-16 ***
## black            5.72654    0.33040   17.332   <2e-16 ***
## hispanic         4.55941    0.51919    8.782   <2e-16 ***
## asian           -0.18726    0.65840   -0.284    0.776
## nativeam        11.60523    0.75787   15.313   <2e-16 ***
## other.race       0.73123    0.93823    0.779    0.436
## male             7.61405    0.28072   27.123   <2e-16 ***
## log(age)         7.11267    0.38426   18.510   <2e-16 ***
## priors           0.49372    0.01542   32.019   <2e-16 ***
## trial_flagTRUE 983.41574   15.62986   62.919   <2e-16 ***
## felony         382.58547    0.78134  489.654   <2e-16 ***
## gross.mis      177.23245    0.74397  238.225   <2e-16 ***
## mis              7.24324    0.30327   23.884   <2e-16 ***
## drug          -154.08645    1.09418 -140.824   <2e-16 ***
## alcohol.dui    -94.03470    0.98219  -95.740   <2e-16 ***
## hunt.fish     -141.66060    2.18805  -64.743   <2e-16 ***
## other.offense -138.03087    0.87246 -158.209   <2e-16 ***
## min_perc         0.33876    0.03589    9.438   <2e-16 ***
## repubvote       -0.01220    0.02499   -0.488    0.626
## lecperc          0.96550    0.04647   20.775   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 218.8 on 2870581 degrees of freedom
##   (2691382 observations deleted due to missingness)
## Multiple R-squared:  0.1613, Adjusted R-squared:  0.1613
```

```
## F-statistic: 2.905e+04 on 19 and 2870581 DF,  p-value: < 2.2e-16
```

```r
#SUR: seemingly unrelated regression
  #same as OLS when covariance matrix is diagonal, or predictors same across equations
sur <- systemfit(list(total_ff_log~black+hispanic+asian+nativeam+other.race+
        male+log(age)+priors+felony+gross.mis+mis+drug+alcohol.dui+hunt.fish+other.offense+
        min_perc+repubvote+lecperc,
                conf_days~black+hispanic+asian+nativeam+other.race+
        male+log(age)+priors+trial_flag+felony+gross.mis+mis+drug+alcohol.dui+hunt.fish+othe
        min_perc+repubvote+lecperc), data=monsanc.short)
```

```
## Warning in systemfit(list(total_ff_log ~ black + hispanic + asian + nativeam + :
## the estimation of systems of equations with unequal numbers of observations has
## not been thoroughly tested yet
```

```r
summary(sur)
```

```
## Warning in nObsEq[i] * nObsEq[j]: NAs produced by integer overflow

## Warning in nObsEq[i] * nObsEq[j]: NAs produced by integer overflow

## Warning in nObsEq[i] * nObsEq[j]: NAs produced by integer overflow

## Warning in nObsEq[i] * nObsEq[j]: NAs produced by integer overflow

##
## systemfit results
## method: OLS
##
##              N       DF          SSR detRCov   OLS-R2 McElroy-R2
## system 5755134 5755095 1.37381e+11 54840.9 0.161261         NA
##
##            N      DF          SSR         MSE        RMSE       R2   Adj R2
## eq1 2884533 2884514 3.29489e+06     1.14227     1.06877 0.117957 0.117952
## eq2 2870601 2870581 1.37377e+11 47856.93489 218.76228 0.161262 0.161256
##
## The covariance matrix of the residuals
##          eq1         eq2
## eq1  1.14634    -4.42575
## eq2 -4.42575 47856.93489
##
## The correlations of the residuals
##            eq1        eq2
## eq1  1.0000000 -0.0188955
## eq2 -0.0188955  1.0000000
##
##
## OLS estimates for 'eq1' (equation 1)
## Model Formula: total_ff_log ~ black + hispanic + asian + nativeam + other.race +
##     male + log(age) + priors + felony + gross.mis + mis + drug +
```

```
##     alcohol.dui + hunt.fish + other.offense + min_perc + repubvote +
##     lecperc
##
##                   Estimate   Std. Error    t value    Pr(>|t|)
## (Intercept)     5.249666657  0.011607682  452.25796 < 2.22e-16 ***
## black          -0.113984773  0.001616223  -70.52541 < 2.22e-16 ***
## hispanic        0.018993451  0.002536874    7.48695 7.0610e-14 ***
## asian           0.012547466  0.003219243    3.89765 9.7135e-05 ***
## nativeam       -0.062835256  0.003704784  -16.96057 < 2.22e-16 ***
## other.race     -0.023608576  0.004586863   -5.14700 2.6471e-07 ***
## male            0.021115614  0.001371081   15.40071 < 2.22e-16 ***
## log(age)       -0.167798954  0.001875480  -89.46987 < 2.22e-16 ***
## priors         -0.007853521  0.000075424 -104.12494 < 2.22e-16 ***
## felony         -0.090944632  0.003820612  -23.80368 < 2.22e-16 ***
## gross.mis       0.555008762  0.003636049  152.64062 < 2.22e-16 ***
## mis             0.510102742  0.001480348  344.58305 < 2.22e-16 ***
## drug            0.403819211  0.005349188   75.49168 < 2.22e-16 ***
## alcohol.dui     0.539693565  0.004800107  112.43365 < 2.22e-16 ***
## hunt.fish      -0.016476542  0.010707965   -1.53872    0.12387
## other.offense   0.261764948  0.004269268   61.31378 < 2.22e-16 ***
## min_perc       -0.023664149  0.000174723 -135.43801 < 2.22e-16 ***
## repubvote       0.001572433  0.000121765   12.91363 < 2.22e-16 ***
## lecperc         0.004663270  0.000226695   20.57068 < 2.22e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.06877 on 2884514 degrees of freedom
## Number of observations: 2884533 Degrees of Freedom: 2884514
## SSR: 3294894.743532 MSE: 1.14227 Root MSE: 1.06877
## Multiple R-Squared: 0.117957 Adjusted R-Squared: 0.117952
##
##
## OLS estimates for 'eq2' (equation 2)
## Model Formula: conf_days ~ black + hispanic + asian + nativeam + other.race +
##     male + log(age) + priors + trial_flag + felony + gross.mis +
##     mis + drug + alcohol.dui + hunt.fish + other.offense + min_perc +
##     repubvote + lecperc
##
##                  Estimate   Std. Error    t value  Pr(>|t|)
## (Intercept)    83.2328554    2.3816634   34.94736  < 2e-16 ***
## black           5.7265366    0.3304025   17.33200  < 2e-16 ***
## hispanic        4.5594149    0.5191923    8.78175  < 2e-16 ***
## asian          -0.1872629    0.6584000   -0.28442  0.77609
## nativeam       11.6052328    0.7578717   15.31293  < 2e-16 ***
## other.race      0.7312340    0.9382339    0.77937  0.43576
## male            7.6140542    0.2807204   27.12326  < 2e-16 ***
## log(age)        7.1126747    0.3842593   18.51009  < 2e-16 ***
## priors          0.4937240    0.0154199   32.01864  < 2e-16 ***
```

```
## trial_flagTRUE   983.4157417    15.6298586    62.91904   < 2e-16 ***
## felony            382.5854731     0.7813392   489.65353   < 2e-16 ***
## gross.mis         177.2324530     0.7439723   238.22455   < 2e-16 ***
## mis                 7.2432388     0.3032718    23.88366   < 2e-16 ***
## drug             -154.0864496     1.0941761  -140.82418   < 2e-16 ***
## alcohol.dui       -94.0347044     0.9821866   -95.74016   < 2e-16 ***
## hunt.fish        -141.6605967     2.1880506   -64.74284   < 2e-16 ***
## other.offense    -138.0308707     0.8724605  -158.20874   < 2e-16 ***
## min_perc            0.3387556     0.0358929     9.43797   < 2e-16 ***
## repubvote          -0.0121953     0.0249871    -0.48806   0.62550
## lecperc             0.9655032     0.0464745    20.77492   < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 218.762279 on 2870581 degrees of freedom
## Number of observations: 2870601 Degrees of Freedom: 2870581
## SSR: 137377208000.093 MSE: 47856.934885 Root MSE: 218.762279
## Multiple R-Squared: 0.161262 Adjusted R-Squared: 0.161256
```

```
#structural equation model
  #2SLS - uses pedicted value of conf days in LFO equation

tsls <- tsls(total_ff_log~black+hispanic+asian+nativeam+other.race+
        male+log(age)+felony+gross.mis+mis+conf_days+drug+alcohol.dui+hunt.fish+other.offens
        min_perc+repubvote+lecperc,
        ~black+hispanic+asian+nativeam+other.race+
        male+log(age)+priors+trial_flag+felony+gross.mis+mis+drug+alcohol.dui+hunt.fish+othe
        min_perc+repubvote+lecperc, data = monsanc.short)

summary(tsls)
```

```
##
##  2SLS Estimates
##
## Model Formula: total_ff_log ~ black + hispanic + asian + nativeam + other.race +
##      male + log(age) + felony + gross.mis + mis + conf_days +
##      drug + alcohol.dui + hunt.fish + other.offense + min_perc +
##      repubvote + lecperc
##
## Instruments: ~black + hispanic + asian + nativeam + other.race + male + log(age) +
##      priors + trial_flag + felony + gross.mis + mis + drug + alcohol.dui +
##      hunt.fish + other.offense + min_perc + repubvote + lecperc
##
## Residuals:
##     Min.  1st Qu.   Median     Mean  3rd Qu.      Max.
##  -6.7585  -0.4168   0.1015   0.0000   0.5351 151.5093
##
##                      Estimate    Std. Error    t value   Pr(>|t|)
```

```
## (Intercept)      5.568983e+00  1.575505e-02  353.47289 < 2.22e-16 ***
## black            -1.171716e-01  2.038468e-03  -57.48022 < 2.22e-16 ***
## hispanic          3.571987e-02  3.109722e-03   11.48652 < 2.22e-16 ***
## asian             2.095320e-02  3.913145e-03    5.35457 8.5768e-08 ***
## nativeam         -3.250359e-02  4.621051e-03   -7.03381 2.0102e-12 ***
## other.race       -2.910777e-02  5.578866e-03   -5.21751 1.8136e-07 ***
## male              3.844084e-02  1.806964e-03   21.27372 < 2.22e-16 ***
## log(age)         -1.536851e-01  2.374836e-03  -64.71398 < 2.22e-16 ***
## felony            1.211583e+00  3.270805e-02   37.04235 < 2.22e-16 ***
## gross.mis         1.149250e+00  1.565113e-02   73.42922 < 2.22e-16 ***
## mis               5.271595e-01  1.915294e-03  275.23694 < 2.22e-16 ***
## conf_days        -3.460134e-03  8.412891e-05  -41.12895 < 2.22e-16 ***
## drug             -1.390720e-01  1.447877e-02   -9.60524 < 2.22e-16 ***
## alcohol.dui       2.265993e-01  9.894497e-03   22.90155 < 2.22e-16 ***
## hunt.fish        -5.073436e-01  1.766894e-02  -28.71386 < 2.22e-16 ***
## other.offense    -2.253476e-01  1.268520e-02  -17.76461 < 2.22e-16 ***
## min_perc         -2.344431e-02  2.158354e-04 -108.62126 < 2.22e-16 ***
## repubvote         1.594466e-03  1.485781e-04   10.73151 < 2.22e-16 ***
## lecperc           7.723260e-03  2.881402e-04   26.80382 < 2.22e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.3007799 on 2870582 degrees of freedom
```