PA2-code and figures

```
library(readr)
library(haven)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##      filter, lag
```

```
## The following objects are masked from 'package:base':
##
##      intersect, setdiff, setequal, union
```

```
library(psych)
library(lavaan)
```

```
## This is lavaan 0.6-9
## lavaan is FREE software! Please report any bugs.
```

```
##
## Attaching package: 'lavaan'
```

```
## The following object is masked from 'package:psych':
##
##      cor2cov
```

```
library(semPlot)
library(texreg)
```

```
## Version:  1.37.5
## Date:     2020-06-17
## Author:   Philip Leifeld (University of Essex)
##
## Consider submitting praise using the praise or praise_interactive functions.
## Please cite the JSS article in your publications -- see citation("texreg").
```

```
data_60 <- read_csv("~/Downloads/data_60.csv")
```

```
## Rows: 635 Columns: 18
```

```
## -- Column specification ------------------------------------------------------
## Delimiter: ","
## dbl (18): id, x1, x2, x3, x4, m1, m2, m3, m4, m5, y1, y2, y3, y4, y5, v15, v...

##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.

View(data_60)

"###Cleaning###"

summary(data_60)

###Remove Unit non response ###
data_60 <- data_60[data_60$id != 145, ]
data_60 <- data_60[data_60$id != 188, ]
data_60 <- data_60[data_60$id != 190, ]
data_60 <- data_60[data_60$id != 258, ]
data_60 <- data_60[data_60$id != 268, ]
data_60 <- data_60[data_60$id != 299, ]
data_60 <- data_60[data_60$id != 314, ]
data_60 <- data_60[data_60$id != 361, ]
data_60 <- data_60[data_60$id != 453, ]
data_60 <- data_60[data_60$id != 549, ]


"### Weird Cases###"

#####Duplicates#####
data60_cl <- add_count(data_60, x1, x2, x3, x4, m1, m2, m3, m4, m5, y1, y2, y3, y4, y5,
                              name = "duplicate") # Name of count variable is "duplicate"

data60_cl<- data60_cl[order(-data60_cl$duplicate),]

data60_cl <- data60_cl[data60_cl$id != 13, ]
data60_cl <- data60_cl[data60_cl$id != 80, ]   #Duplicates cases removed

data60_cl$weird <- 0
data60_cl$weird[data60_cl$duplicate == 2] <- 1   #Add questionable duplicates to weird as 1

#####Straight liners#####
data60_cl$sd <- apply(data60_cl[2:15], 1 , sd)

data60_cl$weird[data60_cl$sd == 0] <- 2
data60_cl <- data60_cl[order(-data60_cl$weird, data60_cl$id),]

####Outliers####
summary(data60_cl)

data60_cl$x1[data60_cl$x1 > 5] <- NA
data60_cl$x2[data60_cl$x2 > 5] <- NA
data60_cl$x3[data60_cl$x3 > 5] <- NA
data60_cl$x4[data60_cl$x4 > 5] <- NA
```

```r
data60_cl$m1[data60_cl$m1 > 5] <- NA
data60_cl$m2[data60_cl$m2 > 5] <- NA
data60_cl$m3[data60_cl$m3 > 5] <- NA
data60_cl$m4[data60_cl$m4 > 5] <- NA
data60_cl$m5[data60_cl$m5 > 5] <- NA
data60_cl$y1[data60_cl$y1 > 5] <- NA
data60_cl$y2[data60_cl$y2 > 5] <- NA
data60_cl$y3[data60_cl$y3 > 5] <- NA
data60_cl$y4[data60_cl$y4 > 5] <- NA
data60_cl$y5[data60_cl$y5 > 5] <- NA

summary(data60_cl) #Recoded incorrect data entries

data60_cl$weird[data60_cl$id == 50] <- 3
data60_cl$weird[data60_cl$id == 370] <- 3
data60_cl$weird[data60_cl$id == 405] <- 3
data60_cl$weird[data60_cl$id == 431] <- 3
data60_cl$weird[data60_cl$id == 450] <- 3
data60_cl$weird[data60_cl$id == 497] <- 3

data60_cl <- data60_cl[order(-data60_cl$weird, data60_cl$id),]

### Check Skewness ###
skew_x1 <- round(skew(data60_cl$x1),2)
skew_x2 <- round(skew(data60_cl$x2),2)
skew_x3 <- round(skew(data60_cl$x3),2)
skew_x4 <- round(skew(data60_cl$x4),2)

skew_m1 <- round(skew(data60_cl$m1),2)
skew_m2 <- round(skew(data60_cl$m2),2)
skew_m3 <- round(skew(data60_cl$m3),2)
skew_m4 <- round(skew(data60_cl$m4),2)
skew_m5 <- round(skew(data60_cl$m5),2)

skew_y1 <- round(skew(data60_cl$y1),2)
skew_y2 <- round(skew(data60_cl$y2),2)
skew_y3 <- round(skew(data60_cl$y3),2)
skew_y4 <- round(skew(data60_cl$y4),2)
skew_y5 <- round(skew(data60_cl$y5),2)

skew_v16 <- round(skew(data60_cl$v16),2)
skew_v17 <- round(skew(data60_cl$v17),2)

data60_sk <- data60_cl

data60_sk$Z_x1 <- scale(data60_sk$x1, center=TRUE, scale=TRUE)
data60_sk$Z_x2 <- scale(data60_sk$x2, center=TRUE, scale=TRUE)
data60_sk$Z_x3 <- scale(data60_sk$x3, center=TRUE, scale=TRUE)
data60_sk$Z_x4 <- scale(data60_sk$x4, center=TRUE, scale=TRUE)

data60_sk$Z_m1 <- scale(data60_sk$m1, center=TRUE, scale=TRUE)
data60_sk$Z_m2 <- scale(data60_sk$m2, center=TRUE, scale=TRUE)
data60_sk$Z_m3 <- scale(data60_sk$m3, center=TRUE, scale=TRUE)
```

```r
data60_sk$Z_m4 <- scale(data60_sk$m4, center=TRUE, scale=TRUE)
data60_sk$Z_m5 <- scale(data60_sk$m5, center=TRUE, scale=TRUE)

data60_sk$Z_y1 <- scale(data60_sk$y1, center=TRUE, scale=TRUE)
data60_sk$Z_y2 <- scale(data60_sk$y2, center=TRUE, scale=TRUE)
data60_sk$Z_y3 <- scale(data60_sk$y3, center=TRUE, scale=TRUE)
data60_sk$Z_y4 <- scale(data60_sk$y4, center=TRUE, scale=TRUE)
data60_sk$Z_y5 <- scale(data60_sk$y5, center=TRUE, scale=TRUE)

summary(data60_sk)


"#### Weird Variables ####"

dfX <- data.frame (first_column  = (data60_cl$x1),
                   second_column = (data60_cl$x2),
                   third_column  = (data60_cl$x3),
                   fourth_column = (data60_cl$x4)
                  )

dfX <- na.exclude(dfX)
names(dfX) = c("x1", "x2", "x3", "x4")
round(cor(dfX), 2)
alpha(dfX)


dfM <- data.frame (first_column  = (data60_cl$m1),
                   second_column = (data60_cl$m2),
                   third_column  = (data60_cl$m3),
                   fourth_column = (data60_cl$m4),
                   fifth_column  = (data60_cl$m5)
)
dfM <- na.exclude(dfM)
names(dfM) = c("m1", "m2", "m3", "m4", "m5")
round(cor(dfM), 2)
alpha(dfM)


dfY <- data.frame (first_column  = (data60_cl$y1),
                   second_column = (data60_cl$y2),
                   third_column  = (data60_cl$y3),
                   fourth_column = (data60_cl$y4),
                   fifth_column  = (data60_cl$y5)
)
dfY <- na.exclude(dfY)
names(dfY) = c("1y", "y2", "y3", "y4", "y5")
round(cor(dfY), 2)
alpha(dfY)

#### Solve M2 ####
data60_cl$m2r[data60_cl$m2 == 1] <- 5
data60_cl$m2r[data60_cl$m2 == 2] <- 4
data60_cl$m2r[data60_cl$m2 == 3] <- 3
```

```r
data60_cl$m2r[data60_cl$m2 == 4] <- 2
data60_cl$m2r[data60_cl$m2 == 5] <- 1

dfM2 <- data.frame (first_column  = (data60_cl$m1),
                    second_column = (data60_cl$m2r),
                    third_column  = (data60_cl$m3),
                    fourth_column = (data60_cl$m4),
                    fifth_column  = (data60_cl$m5)
)
dfM2 <- na.exclude(dfM2)
names(dfM2) = c("m1", "m2r", "m3", "m4", "m5")
round(cor(dfM2), 2)
alpha(dfM2)


"#### Scale Construction ####"
#### Construct new Scales ####
data60_cl$X <- rowMeans(data60_cl[,c(2,3,4,5)], na.rm = TRUE)

data60_cl$M <- rowMeans(data60_cl[,c(6,8,9,10,22)], na.rm = TRUE)

data60_cl$Y <- rowMeans(data60_cl[,c(11,12,13,14,15)], na.rm = TRUE)


"#### Analyses ####"
#### Q1 & Q2 ####
model.1 <- "
M ~ a*X
Y  ~ b*M + cp*X
indirect := a*b
direct   := cp
total    := a*b + cp
"
mediation <- sem(model.1, data = data60_cl, se = "bootstrap", bootstrap=1000)
summary(mediation, ci=T, standardized=T, rsquare=T, fit.measures=F)


#### Q3 & Q4 ####
model.2 <- "
M ~ a*X + d1*v15 +f1*v16
Y  ~ b*M + cp*X + d2*v15 +f2*v16
indirect := a*b + d1 + f1
direct   := cp + d2 + f2
total    := a*b  + d1 + f1 + cp + d2 + f2
"
covariates <- sem(model.2, data = data60_cl, se = "bootstrap", bootstrap=1000)
summary(covariates, ci=T, standardized=T, rsquare=T, fit.measures=F)

#### Q5 ###
model.3 <- "
M ~ a*X + d1*v15 +f1*v16
Y  ~ b*M + cp*X + d2*v15 +f2*v16
v17 ~ g1*M + g2*Y
indirect := a*b + d1 + f1
```

```
direct   := cp  + d2 + f2
total    := a*b  + d1 + f1 + cp + d2 + f2
"
covariate17 <- sem(model.3, data = data60_cl, se = "bootstrap", bootstrap=1000)
summary(covariate17, ci=T, standardized=T, rsquare=T, fit.measures=F)


### Weird ###
data60_weird <- data60_cl
data60_weird1 <- data60_cl
data60_weird3 <- data60_cl
data60_weird <- data60_weird[order(-data60_weird$weird, data60_weird$id),]

data60_weird <- data60_weird[data60_weird$id != 22, ]
data60_weird <- data60_weird[data60_weird$id != 217, ]
data60_weird <- data60_weird[data60_weird$id != 345, ]
data60_weird <- data60_weird[data60_weird$id != 425, ]

data60_weird <- data60_weird[data60_weird$id != 50, ]
data60_weird <- data60_weird[data60_weird$id != 370, ]
data60_weird <- data60_weird[data60_weird$id != 405, ]
data60_weird <- data60_weird[data60_weird$id != 431, ]
data60_weird <- data60_weird[data60_weird$id != 450, ]
data60_weird <- data60_weird[data60_weird$id != 497, ]

data60_weird1 <- data60_weird1[data60_weird1$id != 50, ]
data60_weird1 <- data60_weird1[data60_weird1$id != 370, ]
data60_weird1 <- data60_weird1[data60_weird1$id != 405, ]
data60_weird1 <- data60_weird1[data60_weird1$id != 431, ]
data60_weird1 <- data60_weird1[data60_weird1$id != 450, ]
data60_weird1 <- data60_weird1[data60_weird1$id != 497, ]

data60_weird3 <- data60_weird3[data60_weird3$id != 22, ]
data60_weird3 <- data60_weird3[data60_weird3$id != 217, ]
data60_weird3 <- data60_weird3[data60_weird3$id != 345, ]
data60_weird3 <- data60_weird3[data60_weird3$id != 425, ]


mediation.W <- sem(model.1, data = data60_weird, se = "bootstrap", bootstrap=1000)
summary(mediation.W, ci=T, standardized=T, rsquare=T, fit.measures=F)


covariates.W <- sem(model.2, data = data60_weird, se = "bootstrap", bootstrap=1000)
summary(covariates.W, ci=T, standardized=T, rsquare=T, fit.measures=F)


covariate17.W <- sem(model.3, data = data60_weird, se = "bootstrap", bootstrap=1000)
summary(covariate17.W, ci=T, standardized=T, rsquare=T, fit.measures=F)


### Omitted Var Bias ###

screenreg(list(mediation, mediation.W),      # Names of the R-objects from above
```

```
          custom.model.name =        # Give new, descriptive names to the models
            c("Model 1: Normal",
              "Model 2: Adjusted for Weird"),
          single.row = TRUE, digits = 3)




screenreg(list(covariates, covariates.W),      # Names of the R-objects from above
          custom.model.name =        # Give new, descriptive names to the models
            c("Model 1: Normal",
              "Model 2: Adjusted for Weird"),
          single.row = TRUE, digits = 3)




screenreg(list(covariate17, covariate17.W),      # Names of the R-objects from above
          custom.model.name =        # Give new, descriptive names to the models
            c("Model 1: Normal",
              "Model 2: Adjusted for Weird"),
          single.row = TRUE, digits = 3)


### Omitted Var Bias 2 ###

mediation.W1 <- sem(model.1, data = data60_weird1, se = "bootstrap", bootstrap=1000)
summary(mediation.W1, ci=T, standardized=T, rsquare=T, fit.measures=F)


covariates.W1 <- sem(model.2, data = data60_weird1, se = "bootstrap", bootstrap=1000)
summary(covariates.W1, ci=T, standardized=T, rsquare=T, fit.measures=F)


covariate17.W1 <- sem(model.3, data = data60_weird1, se = "bootstrap", bootstrap=1000)
summary(covariate17.W1, ci=T, standardized=T, rsquare=T, fit.measures=F)



screenreg(list(mediation.W, mediation.W1),      # Names of the R-objects from above
          custom.model.name =        # Give new, descriptive names to the models
            c("Model 1: Adjusted for Weird",
              "Model 2: Potential duplicates removed"),
          single.row = TRUE, digits = 3)



screenreg(list(covariates.W, covariates.W1),      # Names of the R-objects from above
          custom.model.name =        # Give new, descriptive names to the models
            c("Model 1: Adjusted for Weird",
              "Model 2: Potential duplicates removed"),
          single.row = TRUE, digits = 3)
```

```
screenreg(list(covariate17.W, covariate17.W1),      # Names of the R-objects from above
          custom.model.name =         # Give new, descriptive names to the models
            c("Model 1: Adjusted for Weird",
              "Model 2: Potential duplicates removed"),
          single.row = TRUE, digits = 3)


################################################################

mediation.W3 <- sem(model.1, data = data60_weird3, se = "bootstrap", bootstrap=1000)
summary(mediation.W3, ci=T, standardized=T, rsquare=T, fit.measures=F)


covariates.W3 <- sem(model.2, data = data60_weird3, se = "bootstrap", bootstrap=1000)
summary(covariates.W3, ci=T, standardized=T, rsquare=T, fit.measures=F)


covariate17.W3 <- sem(model.3, data = data60_weird3, se = "bootstrap", bootstrap=1000)
summary(covariate17.W3, ci=T, standardized=T, rsquare=T, fit.measures=F)


screenreg(list(mediation.W, mediation.W3),      # Names of the R-objects from above
          custom.model.name =         # Give new, descriptive names to the models
            c("Model 1: All Weird removed",
              "Model 2: Item Non Response removed"),
          single.row = TRUE, digits = 3)


screenreg(list(covariates.W, covariates.W3),      # Names of the R-objects from above
          custom.model.name =         # Give new, descriptive names to the models
            c("Model 1: All Weird removed",
              "Model 2: Item Non Response removed"),
          single.row = TRUE, digits = 3)


screenreg(list(covariate17.W, covariate17.W3),      # Names of the R-objects from above
          custom.model.name =         # Give new, descriptive names to the models
            c("Model 1: All Weird removed",
              "Model 2: Item Non Response removed"),
          single.row = TRUE, digits = 3)
```