Data from National Youth Survey

Each year, when participants where ages 11, 12, 13, 14, and 15, they filled out a nine-item instrument designed to assess their tolerance of deviant behavior.

Using a four-point scale, where 1 = very wrong and 4 = not wrong at all, they indicated whether it was wrong for someone their age to (a) cheat on tests, (b) purposely destroy the property of others (c) use marijuana (d) steal something worth less than five dollars € hit or threaten someone without reason (f) use alcohol, (g) break into a building or vehicle to steal (h) sell hard drugs, or (i) steal something worth more than 50 dollars. At each occasion, the outcome (TOL) is computed as the respondent’s average across nine responses. Respondents also recorded their sex and exposure, a self-reported rating of deviant behavior at age 11. Exposure is measured by the respondent’s rating of involvement of close friends’ in the nine deviant behaviors, and is also an average of those ratings.

title1 "Person-level Tolerance Data";

**data** tolerance;

input OBS ID Tol11 Tol12 Tol13 Tol14 Tol15 Male Exposure;

datalines;

1 9 2.23 1.79 1.90 2.12 2.66 0 1.54

2 45 1.12 1.45 1.45 1.45 1.99 1 1.16

3 268 1.45 1.34 1.99 1.79 1.34 1 0.90

4 314 1.22 1.22 1.55 1.12 1.12 0 0.81

5 442 1.45 1.99 1.45 1.67 1.90 0 1.13

6 514 1.34 1.67 2.23 2.12 2.44 1 0.90

7 569 1.79 1.90 1.90 1.99 1.99 0 1.99

8 624 1.12 1.12 1.22 1.12 1.22 1 0.98

9 723 1.22 1.34 1.12 1.00 1.12 0 0.81

10 918 1.00 1.00 1.22 1.99 1.22 0 1.21

11 949 1.99 1.55 1.12 1.45 1.55 1 0.93

12 978 1.22 1.34 2.12 3.46 3.32 1 1.59

13 1105 1.34 1.90 1.99 1.90 2.12 1 1.38

14 1542 1.22 1.22 1.99 1.79 2.12 0 1.44

15 1552 1.00 1.12 2.23 1.55 1.55 0 1.04

16 1653 1.11 1.11 1.34 1.55 2.12 0 1.25

run;

**proc** **print** data=tolerance; **run**;

\*Creating a person-period data set from a balanced person-level data set;

**data** tolerance\_pp;

set tolerance;

array Atol [**11**:**15**] tol11-tol15;

do age=**11** to **15**;

tol = Atol[age];

time = age - **11**;

output;

end;

keep id age time tol male exposure;

**run**;

title1 "Person-Period data set";

**proc** **print** data=tolerance\_pp;

var id age tol male exposure;

**run**;

title1 'Estimated bivariate correlations among tolerance scores';

options nocenter nodate nolabel;

**proc** **corr** data='c:\alda\tolerance' nosimple noprob;

var tol11-tol15;

**run**;

\*Panel Plot – Change over time;

**proc** **sgpanel** data = tolerance\_pp;

panelby id /columns=**4** rows= **4**;

scatter y = tol x = age;

**run**;

\*Panel Plot with curve estimate;

**proc** **sgpanel** data = tolerance\_pp (rename=(tol=tolerance));

panelby id /columns=**4** rows= **4**;

pbspline y = tolerance x = age;

**run**;

\*Fitting separate exploratory regression models;

**proc** **sort** data = tolerance\_pp;

by male exposure id;

**run**;

**proc** **reg** data = tolerance\_pp

outest = expl (drop = \_MODEL\_ \_in\_ \_p\_ \_edf\_) outseb rsquare noprint;

by male exposure id;

model tol = time;

**run**;

**proc** **sort** data = expl;

by id \_type\_;

**run**;

**data** expl2 (rename = (a0=intercept a1 = time));

set expl;

by id;

r2 = lag(\_rsq\_);

retain a0 se\_intercept a1 se\_time res\_var r2;

if first.id then do;

a0 = intercept;

a1 = time;

end;

if \_type\_ ="SEB" then do;

se\_intercept = intercept;

se\_time = time;

res\_var = \_rmse\_\*\***2**;

end;

if last.id;

keep id a0 se\_intercept a1 se\_time res\_var r2 male exposure;

**run**;

**proc** **print** data = expl2 noobs;

var id intercept se\_intercept time se\_time res\_var r2 male exposure;

format \_numeric\_ **4.2**;

format id **4.0**;

**run**;

\*Panel plot of OLS Fits;

**proc** **sgpanel** data = tolerance\_pp (rename=(tol=tolerance));

panelby id /columns=**4** rows= **4**;

reg y = tolerance x = age;

**run**;

**quit**;

\*Trajectories for subjects across time;

**proc** **sgplot** data=tolerance\_pp (rename=(tol=tolerance)) noautolegend ;

\* spaghetti plot;

yaxis min = **0** max = **4**;

pbspline x=age y=tolerance

/ group = id nomarkers LINEATTRS = (COLOR= gray PATTERN = **1** THICKNESS = **1**) ;

\* overall spline;

pbspline x=age y=tolerance

/ nomarkers LINEATTRS = (COLOR= red PATTERN = **1** THICKNESS = **3**) ;

**run**;

**quit**;

**proc** **sgplot** data=tolerance\_pp (rename=(tol=tolerance)) noautolegend ;

\* spaghetti plot;

yaxis min = **0** max = **4**;

reg x=age y=tolerance

/ group = id nomarkers LINEATTRS = (COLOR= gray PATTERN = **1** THICKNESS = **1**) ;

\* overall spline;

reg x=age y=tolerance

/ nomarkers LINEATTRS = (COLOR= red PATTERN = **1** THICKNESS = **3**) ;

**run**;

**quit**;

\*Descriptive statistics obtained by fitting separate OLS Curves;

**proc** **corr** data = expl2;

var intercept time;

**run**;

\*Change over time by sex and exposure;

**proc** **sgpanel** data=tolerance\_pp (rename=(tol=tolerance)) noautolegend ;

panelby male;

reg x=age y=tolerance

/ group = id nomarkers LINEATTRS = (COLOR= gray PATTERN = **1** THICKNESS = **1**) ;

\* overall spline;

reg x=age y=tolerance

/ nomarkers LINEATTRS = (COLOR= red PATTERN = **1** THICKNESS = **3**) ;

**run**;

**proc** **means** data = tolerance\_pp median;

var exposure;

output out = t median = m;

**run**;

**data** \_null\_;

set t;

call symput('exp', m);

**run**;

**proc** **format**;

value exp **0** = "Low exposure"

**1** = "High exposure";

**run**;

**data** to\_exp;

set tolerance\_pp;

if exposure < &exp then exp\_cat = **0**;

else exp\_cat = **1**;

format exp\_cat exp.;

rename tol = tolerance;

**run**;

**proc** **sgpanel** data=to\_exp noautolegend ;

panelby exp\_cat;

reg x=age y=tolerance

/ group = id nomarkers LINEATTRS = (COLOR= gray PATTERN = **1** THICKNESS = **1**) ;

\* overall spline;

reg x=age y=tolerance

/ nomarkers LINEATTRS = (COLOR= red PATTERN = **1** THICKNESS = **3**) ;

**run**;

**quit**;

\*Examine relationship between OLS parameters and predictors;

ods output PearsonCorr = pcorr;

**proc** **corr** data = expl2 nosimple noprob;

var intercept time;

with male exposure;

**run**;

**data** \_null\_;

set pcorr;

if \_n\_ = **1** then do;

call symput('male\_int', put(intercept, **4.2**));

call symput('male\_time', put(time, **4.2**));

end;

if \_n\_ = **2** then do;

call symput('exp\_int', put(intercept, **4.2**));

call symput('exp\_time', put(time, **4.2**));

end;

**run**;

title "Examining relationship";

**proc** **template**;

define statgraph fourPlots;

begingraph;

layout gridded / columns=**2** rows=**2** ;

layout overlay / yaxisopts=(linearopts=(viewmin=**.5** viewmax= **2.5**))

xaxisopts=(linearopts=(viewmin=-**.5** viewmax= **1.5**

tickvaluelist=(**0** **1**))) autoalign=(topleft topright);

scatterplot x=male y=intercept;

layout gridded / autoalign=(BottomRight) border = false ;

entry halign=left "r = &male\_int" ;

endlayout;

endlayout;

layout overlay / yaxisopts=(linearopts=(viewmin=**.5** viewmax= **2.5**))

xaxisopts=(linearopts=(viewmin=**0** viewmax= **2.5**

tickvaluelist=(**0** **1** **2**)));

scatterplot x=exposure y=intercept;

layout gridded / autoalign=(BottomRight) border = false ;

entry halign=left "r = &exp\_int" ;

endlayout;

endlayout;

layout overlay / yaxisopts=(linearopts=(viewmin=**0** viewmax= **.8**))

xaxisopts=(linearopts=(viewmin=-**0.5** viewmax= **1.5**

tickvaluelist=(**0** **1**)));

scatterplot x=male y=time;

layout gridded / autoalign=(BottomRight) border = false ;

entry halign=left "r = &male\_time" ;

endlayout;

endlayout;

layout overlay / yaxisopts=(linearopts=(viewmin=**0** viewmax= **.8**))

xaxisopts=(linearopts=(viewmin=**0** viewmax= **2.5**

tickvaluelist=(**0** **1** **2**)));

scatterplot x=exposure y=time;

layout gridded / autoalign=(BottomRight) border = false ;

entry halign=left "r = &exp\_time" ;

endlayout;

endlayout;

endlayout;

endgraph;

end;

**run**;

**proc** **sgrender** data=expl2 template=fourPlots;

**run**;