VISUALIZATIONS AND MPLUS SYNTAX FOR SURVIVAL MODELING

[EXAMPLE DATASET 1 2](#_Toc515363049)

[INPUT EX 1a: Discrete-Time Survival 3](#_Toc515363050)

[OUTPUT EX 1a 4](#_Toc515363051)

[INPUT EX 1b: Continuous-Time Survival 7](#_Toc515363052)

[OUTPUT EX 1b 8](#_Toc515363053)

[INPUT EX 2: Left-Censoring 9](#_Toc515363054)

[OUTPUT EX 2 11](#_Toc515363055)

[EXAMPLE DATASET 2 12](#_Toc515363056)

[INPUT EX 3a: Time-varying Predictors 13](#_Toc515363057)

[OUTPUT EX 3a 14](#_Toc515363058)

[INPUT EX 3b: Time-varying Effects 16](#_Toc515363059)

[OUTPUT EX 3b 17](#_Toc515363060)

[INPUT EX 4a: Test of Proportional Hazards assumption - Independent 19](#_Toc515363061)

[OUTPUT EX 4a 20](#_Toc515363062)

[INPUT EX 4b: Test of Proportional Hazards assumption - Linear 24](#_Toc515363063)

[OUTPUT EX 4b: 25](#_Toc515363064)

[INPUT EX 5a: Mediation 26](#_Toc515363065)

[OUTPUT EX 5a 30](#_Toc515363066)

[INPUT EX 5b: Time-invariant Mediation 34](#_Toc515363067)

[OUTPUT EX 5b 36](#_Toc515363068)

[INPUT EX 5c: Moderated Mediation 39](#_Toc515363069)

[OUTPUT EX 5c 42](#_Toc515363070)

[INPUT EX 6a: Multiple Survival Curves 45](#_Toc515363071)

[OUTPUT EX 6a 47](#_Toc515363072)

[INPUT EX 6b: Test of Proportional Hazards Assumption, Revisited 53](#_Toc515363073)

[OUTPUT EX 6b 55](#_Toc515363074)

[INPUT EX 6c: Quadratic Proportional Hazards Violation 57](#_Toc515363075)

[OUTPUT EX 6c 59](#_Toc515363076)

[References: 62](#_Toc515363077)

# EXAMPLE DATASET 1

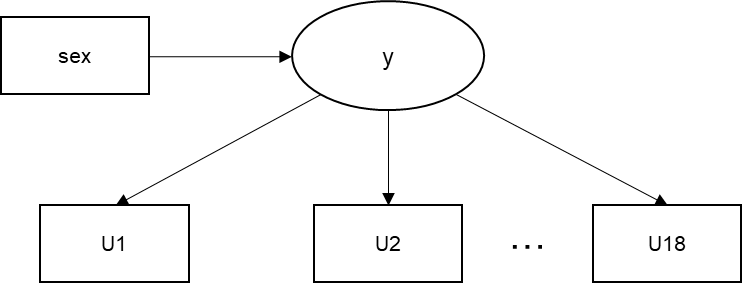


When using Mplus, a wide-style dataset is typically used. For discrete-time data, each time point at which the event may occur is assigned its own column. For use in Mplus, variable names are not included. ‘1’ indicates the event has occurred, while ‘0’ indicates the event has not yet occurred. In the case here, the data was transformed prior to use in Mplus in the same style that DATA SURVIVAL would transform the data.

That is, all cases of data after event occurrence or after the last observation of non-initiation has occurred, are assigned the missing data value of 999.

In our case, the variables represent half-age progressions, where MJ1 identifies event occurrence at age 11.5 years.

# INPUT EX 1a: Discrete-Time Survival



TITLE: EX 1a: DISCRETE-TIME SURVIVAL MODEL

DATA:

FILE IS SurvivalExampleData.dat;

VARIABLE:

!Mplus reads data files by order of columns, not names. To read a data file in Mplus,

!Remove any headers containing variable names.

!Then make a list of variable names by COLUMN ORDER. This is easy to mess up.

!Mplus will read as many columns as you supply variable names.

NAMES ARE ID

MJ1 - MJ26

DRK1 - DRK26

puff1 - puff26

sex !Categorical variable with values 0 and 1

hisp !Categorical variable with values 0 and 1

white !Categorical variable with values 0 and 1

cohort !Categorical variable with values 1-5

;

!All survival curve indicators must have at least some cases with values ‘0’ and ‘1’.

!If this is not the case, the data point may be omitted from analyses.

!But note that when plotting changes over time, or graphing, the spacing between

!data points will be uneven.

!In this case, this is why we use mj1-mj18, instead of mj1-mj26.

!In USEVARIABLES, always list source data before newly created variables.

USEVARIABLES = sex U1-U18;

Missing = all(999);

CATEGORICAL = U1-U18;

DSURVIVAL = U1-U18;

!DATA SURVIVAL is the command that Mplus uses to format data in order to be ready for

!survival analysis. This command automatically censors any observations of data after

!an event occurs, and checks the data for any additional cases of values not equal to

!0 and 1 (and will tell you if there are violations). See Mplus User’s Guide, Version

!8, page 590.

!Researchers can do these transformations BY HAND, and use just the original data instead.

DATA SURVIVAL:

NAMES = MJ1-MJ18; !Source data - we use raw onset data on marijuana use.

CUTPOINT = 0; !Tells Mplus to discriminate between 0 and 1 for event

BINARY = U1-U18; !New transformed data for Model statement

ANALYSIS:

process=4;

!In many cases where data are spaced by age, there will be indicators with no

!overlap. To account for this, the command below tells Mplus to

!ignore warnings regarding "missing covariance coverage".

coverage=0;

ESTIMATOR = MLR;

MODEL:

y by U1-U18@1;

!By constraining y to have no residual variance, we can interpret the thresholds of the

!indicators as time-specific logit hazard values, after controlling for predictors that

!affect y.

y on sex;

y@0;

# OUTPUT EX 1a

*Results from the Discrete-Time Survival Analysis, graphed using Excel to calculate cumulative hazard for girls (thresholds only) and boys (thresholds + effect of sex).*

EX 1a: DISCRETE-TIME SURVIVAL MODEL, PH ASSUMED

SUMMARY OF ANALYSIS

Number of groups 1

Number of observations 1023

Number of dependent variables 18

Number of independent variables 1

Number of continuous latent variables 1

THE MODEL ESTIMATION TERMINATED NORMALLY

MODEL FIT INFORMATION

Model fit statistics can be used to calculate LRTs, in order to examine whether models improve enough to justify increases in #parameters. Note the MLR scaling factor and see https://stats.idre.ucla.edu/mplus/faq/how-can-i-compute-a-chi-square-test-for-nested-models-with-the-mlr-or-mlm-estimators/.

Number of Free Parameters 19

Loglikelihood

H0 Value -1717.027

H0 Scaling Correction Factor 1.0002

for MLR

Information Criteria

Akaike (AIC) 3472.054

Bayesian (BIC) 3565.733

Sample-Size Adjusted BIC 3505.387

(n\* = (n + 2) / 24)

MODEL RESULTS

Two-Tailed

Estimate S.E. Est./S.E. P-Value

Y BY

U1 1.000 0.000 999.000 999.000

U2 1.000 0.000 999.000 999.000

U3 1.000 0.000 999.000 999.000

U4 1.000 0.000 999.000 999.000

U5 1.000 0.000 999.000 999.000

U6 1.000 0.000 999.000 999.000

U7 1.000 0.000 999.000 999.000

U8 1.000 0.000 999.000 999.000

U9 1.000 0.000 999.000 999.000

U10 1.000 0.000 999.000 999.000

U11 1.000 0.000 999.000 999.000

U12 1.000 0.000 999.000 999.000

U13 1.000 0.000 999.000 999.000

U14 1.000 0.000 999.000 999.000

U15 1.000 0.000 999.000 999.000

U16 1.000 0.000 999.000 999.000

U17 1.000 0.000 999.000 999.000

U18 1.000 0.000 999.000 999.000

Y ON

SEX -0.225 0.102 -2.201 0.028

The effect of sex on time-invariant proportional hazard is significant. Since sex is coded 1 for boys and 0 for girls, and the logit effect is negative, we conclude that girls are at greater overall risk for initiation of marijuana use in this sample.

The odds of marijuana initiation at each time point, given that a participant is male as compared to female, are exp(-0.225)= .80.

Thresholds

U1$1 6.828 0.999 6.835 0.000

U2$1 5.437 0.502 10.826 0.000

U3$1 4.326 0.294 14.728 0.000

U4$1 4.493 0.320 14.052 0.000

U5$1 3.879 0.240 16.134 0.000

U6$1 3.551 0.209 17.002 0.000

U7$1 3.213 0.186 17.249 0.000

U8$1 2.945 0.170 17.315 0.000

U9$1 3.219 0.200 16.098 0.000

U10$1 3.126 0.194 16.104 0.000

U11$1 2.965 0.186 15.943 0.000

U12$1 2.672 0.172 15.500 0.000

U13$1 2.945 0.203 14.508 0.000

U14$1 2.713 0.191 14.237 0.000

U15$1 2.587 0.187 13.816 0.000

U16$1 2.692 0.206 13.037 0.000

U17$1 2.375 0.221 10.762 0.000

U18$1 3.227 0.590 5.471 0.000

Thresholds are time-specific logit hazards that cumulatively describe the hazard of onset over the course of the time period (when all predictors are at zero, so, for girls).

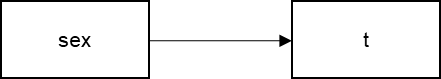
Users can take these thresholds and graph the onset rate for boys and girls. See Singer & Willett, 2003, p.386-392.

Note that Mplus handles dichotomous variables slightly differently than most other statistical software. Since Mplus predicts probability of being in the first value of a categorical variable (relative to being in the last), these tresholds are flipped in direction. In order to calculate the hazard of event occurrence, we flip the sign of these estimates (-1\*U1$1=-6.828 for the logit hazard of onset at time point U1).

Residual Variances

Y 0.000 0.000 999.000 999.000

# INPUT EX 1b: Continuous-Time Survival



TITLE: EX 1a: CONTINUOUS-TIME SURVIVAL MODEL, PH ASSUMED

!Based on Mplus user's guide, example 6.20.

DATA:

FILE IS SurvivalExampleData.dat;

VARIABLE:

!To illustrate parallels between continuous-time and discrete-time survival

!modeling, we will re-code data into continuous-time outcomes here.

!Note that these variables are already prepared for discrete-time survival models.

NAMES ARE ID

MJ1 - MJ26

DRK1 - DRK26

puff1 - puff26

sex

hisp

white

cohort

;

USEVARIABLES = sex t tc;

survival = t;

timecensored = tc (0=NOT 1=RIGHT);

Missing = all(999);

DEFINE:

!Re-coding data into continuous-time data. Mplus is not an efficient way to do this.

if MJ1 NE 999 THEN t=1;

if MJ1 EQ 1 THEN tc=0;

if MJ1 EQ 0 THEN tc=1;

if MJ2 NE 999 THEN t=2;

if MJ2 EQ 1 THEN tc=0;

if MJ2 EQ 0 THEN tc=1;

!...Repeat for all variables. Abbreviated here.

if MJ17 NE 999 THEN t=18;

if MJ17 EQ 1 THEN tc=0;

if MJ17 EQ 0 THEN tc=1;

if MJ18 NE 999 THEN t=19;

if MJ18 EQ 1 THEN tc=0;

if MJ18 EQ 0 THEN tc=1;

ANALYSIS:

process=4;

ESTIMATOR = MLR;

MODEL:

!Because we have not ultimately specified our continuous data on a scale differently than

!the one we specified for the discrete-time survival curve, we find that the effect of

!sex is identical to that from example 1a.

t on sex;

# OUTPUT EX 1b

EX 1b: CONTINUOUS-TIME SURVIVAL MODEL

SUMMARY OF ANALYSIS

Number of groups 1

Number of observations 1023

Number of dependent variables 1

Number of independent variables 1

Number of continuous latent variables 0

Observed dependent variables

Time-to-event (survival)

Non-parametric

T

Observed independent variables

SEX

Variables with special functions

Time-censoring variables

TC

THE MODEL ESTIMATION TERMINATED NORMALLY

MODEL FIT INFORMATION

Number of Free Parameters 1

!The LRT cannot be used to compare discrete-time with continuous-time models. They are not nested.

Loglikelihood

H0 Value -1745.163

H0 Scaling Correction Factor 0.9539

for MLR

Information Criteria

!AIC and BIC also require identical outcome data in order for fair comparisons to occur. Therefore, this AIC and BIC alsocannot be compared to example 1a.

Akaike (AIC) 3492.326

Bayesian (BIC) 3497.256

Sample-Size Adjusted BIC 3494.080

(n\* = (n + 2) / 24)

MODEL RESULTS

Two-Tailed

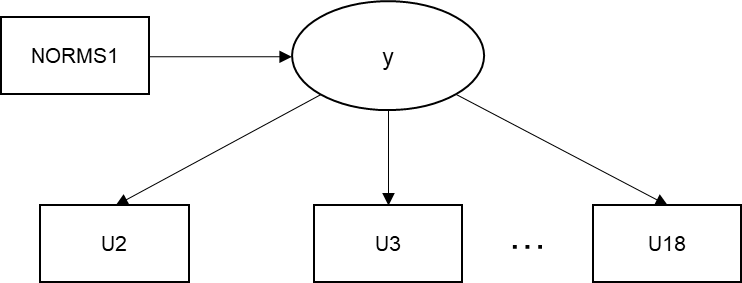
Estimate S.E. Est./S.E. P-Value

T ON

SEX -0.215 0.098 -2.200 0.028

!We find a nearly identical effect of sex on time to survival.

# INPUT EX 2: Left-Censoring



TITLE: EX 2: LEFT CENSORING, BY BASELINE TIME POINT

DATA:

FILE IS SurvivalExampleData2.dat;

VARIABLE:

NAMES ARE ID

MJ1 - MJ26

DRK1 - DRK26

puff1 - puff26

sex

hisp

white

cohort

T1MJ

POS1-POS17 !Positive Expectancies (continuous)

NEG1-NEG17 !Negative Expectancies (continuous)

AVAIL1-AVAIL17 !Alcohol Availability (Dichotomous)

Norms1- Norms17 !Perceived Norms (continuous)

;

!The USEOBSERVATIONS command can be used to exclude assessments that were concurrent to predictor !assessments, if violation of temporal ordering is to be avoided.

!So, if we were interested in a purely prospective effect for Norms (Peer Norms as a prospective !predictor of initiation), given that it might itself change over time as a consequence of !initiation (temporal order is of interest to the researcher), we would exclude observations that !came from the time point at which Norms1 was assessed (baseline),

!and because that time point corresponded to different ages for different participants, we use a !pre-coded variable for this purpose.

USEOBSERVATIONS= T1MJ NE 1;

!Because observations from baseline were excluded, variable 'u1' no longer has more than 1 !category (no remaining participants onset at this time point), hence it had to be excluded.

USEVARIABLES = Norms1 U2-U18;

Missing = all(999);

CATEGORICAL = U2-U18;

DSURVIVAL = U2-U18;

DATA SURVIVAL:

NAMES = MJ2-MJ18;

CUTPOINT = 0;

BINARY = U2-U18;

ANALYSIS:

process=4;

coverage=0;

ESTIMATOR = MLR;

MODEL:

y by U2-U18@1;

y on Norms1;

y@0;

# OUTPUT EX 2

SUMMARY OF ANALYSIS

Number of groups 1

!The reduced number of observations reflects those excluded due to concurrent marijuana initiation.

Number of observations 957

Number of dependent variables 17

Number of independent variables 1

Number of continuous latent variables 1

MODEL FIT INFORMATION

Number of Free Parameters 18

!Models with a different set of observations are not nested and cannot be compared to previous.

Loglikelihood

H0 Value -1438.580

H0 Scaling Correction Factor 1.0029

for MLR

Information Criteria

Akaike (AIC) 2913.160

Bayesian (BIC) 3000.538

Sample-Size Adjusted BIC 2943.371

(n\* = (n + 2) / 24)

MODEL RESULTS

Two-Tailed

Estimate S.E. Est./S.E. P-Value

Y ON

NORMS1 0.159 0.061 2.605 0.009

Thresholds

U2$1 6.937 1.003 6.919 0.000

U3$1 5.545 0.504 11.003 0.000

U4$1 5.313 0.450 11.810 0.000

U5$1 4.708 0.335 14.054 0.000

U6$1 4.037 0.245 16.450 0.000

U7$1 4.128 0.263 15.713 0.000

U8$1 3.374 0.187 18.049 0.000

U9$1 3.542 0.207 17.149 0.000

U10$1 3.367 0.200 16.846 0.000

U11$1 3.166 0.186 17.048 0.000

U12$1 2.841 0.169 16.844 0.000

U13$1 3.153 0.203 15.502 0.000

U14$1 2.884 0.188 15.302 0.000

U15$1 2.755 0.187 14.711 0.000

U16$1 2.861 0.204 14.025 0.000

U17$1 2.533 0.219 11.547 0.000

U18$1 3.391 0.591 5.739 0.000

Residual Variances

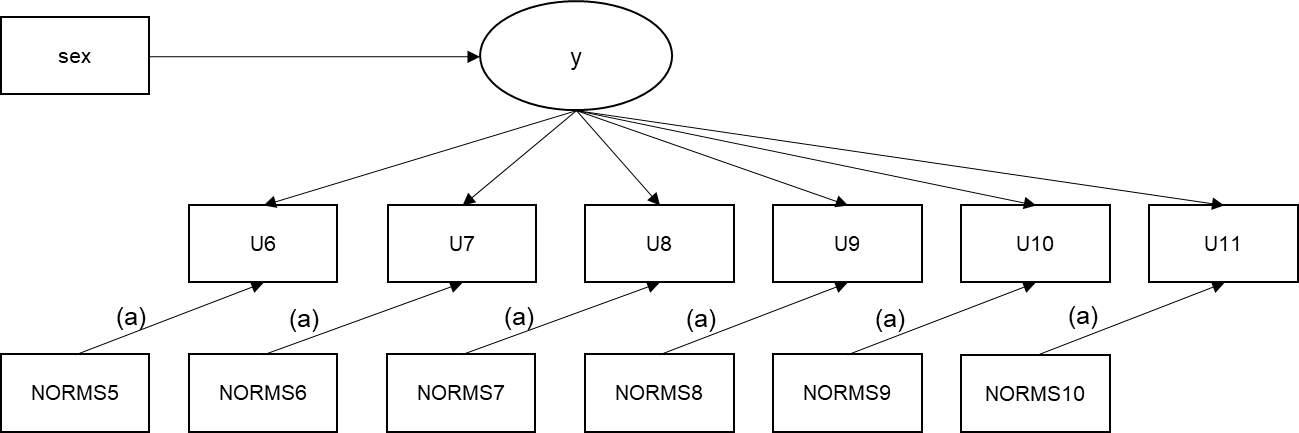
Y 0.000 0.000 999.000 999.000

# EXAMPLE DATASET 2



Time-varying predictors (here: lagged-1 predictor NORMS#) are separate variables here, with numbers for corresponding time frames to MJ time frames. A single column has values for the time-invariant predictor sex.

# INPUT EX 3a: Time-varying Predictors



*The parameter denotation (a) indicates that the effect of all pathways with that denotation are assumed to be identical.*

TITLE: EX 3a: TIME-VARYING PREDICTORS, LAGGED (TIME-INVARIANT) EFFECT, CONTROLLING FOR TIME-INVARIANT COVARIATE

DATA:

FILE IS SurvivalExampleData4.dat;

VARIABLE:

NAMES ARE ID

MJ1 - MJ26

DRK1 - DRK26

puff1 - puff26

sex

hisp

white

cohort

T1MJ

POS1-POS17 !Positive Expectancies (continuous)

NEG1-NEG17 !Negative Expectancies (continuous)

AVAIL1-AVAIL17 !Alcohol Availability (Dichotomous)

NORMS1-NORMS17 !Perceived Norms (continuous)

;

USEOBSERVATIONS= (MJ2 NE 1) AND (MJ3 NE 1) AND (MJ4 NE 1) AND (MJ5 NE 1);

USEVARIABLES = sex NORMS5-NORMS10 U6-U11;

Missing = all(999);

CATEGORICAL = U6-U11;

DSURVIVAL = U6-U11;

DATA SURVIVAL:

NAMES = MJ6-MJ11;

CUTPOINT = 0;

BINARY = U6-U11;

ANALYSIS:

process=4;

coverage=0;

ESTIMATOR = MLR;

!Accounting for missing on predictors necessitates integration.

ALGORITHM = integration;

integration = Montecarlo;

MODEL:

!We assume that time-varying NORMS5-NORMS10, assessed at time points corresponding to MJ5-MJ10,

!have lagged effects on initiation, after controlling for the overall (time-invariant) effect of

!baseline predictor 'sex'.

y by U6-U11@1;

y on sex;

y@0;

!In input syntax, giving parameters the same label constrains them to be equal to each other. This can also be achieved with the MODEL CONSTRAINT command as referenced above.

U6 on NORMS5 (a);

U7 on NORMS6 (a);

U8 on NORMS7 (a);

U9 on NORMS8 (a);

U10 on NORMS9 (a);

U11 on NORMS10 (a);

# OUTPUT EX 3a

EX 3a: TIME-VARYING PREDICTORS, LAGGED (TIME-INDEPENDENT) EFFECT, CONTROLLING FOR TIME-INVARIANT C

SUMMARY OF ANALYSIS

Number of groups 1

Number of observations 979

!Because we deal with missing data on predictors, all observations are retained (save those excluded by USEOBSERVATIONS statements.

!However, note assumptions about the cause of missing data on predictors.

Number of dependent variables 6

Number of independent variables 7

Number of continuous latent variables 1

THE MODEL ESTIMATION TERMINATED NORMALLY

MODEL FIT INFORMATION

Number of Free Parameters 35

Loglikelihood

H0 Value -12292.757

H0 Scaling Correction Factor 1.4878

for MLR

Information Criteria

Akaike (AIC) 24655.515

Bayesian (BIC) 24826.544

Sample-Size Adjusted BIC 24715.383

(n\* = (n + 2) / 24)

MODEL RESULTS

Two-Tailed

Estimate S.E. Est./S.E. P-Value

Y BY

U6 1.000 0.000 999.000 999.000

U7 1.000 0.000 999.000 999.000

U8 1.000 0.000 999.000 999.000

U9 1.000 0.000 999.000 999.000

U10 1.000 0.000 999.000 999.000

U11 1.000 0.000 999.000 999.000

Y ON

SEX -0.319 0.154 -2.074 0.038

!The lagged effects of U on NORMS are adjusted for the time-invariant effect of y on sex.

U6 ON

NORMS5 0.113 0.017 6.636 0.000

U7 ON

NORMS6 0.113 0.017 6.636 0.000

U8 ON

NORMS7 0.113 0.017 6.636 0.000

U9 ON

NORMS8 0.113 0.017 6.636 0.000

U10 ON

NORMS9 0.113 0.017 6.636 0.000

U11 ON

NORMS10 0.113 0.017 6.636 0.000

Thresholds

U6$1 3.649 0.210 17.381 0.000

U7$1 3.361 0.193 17.426 0.000

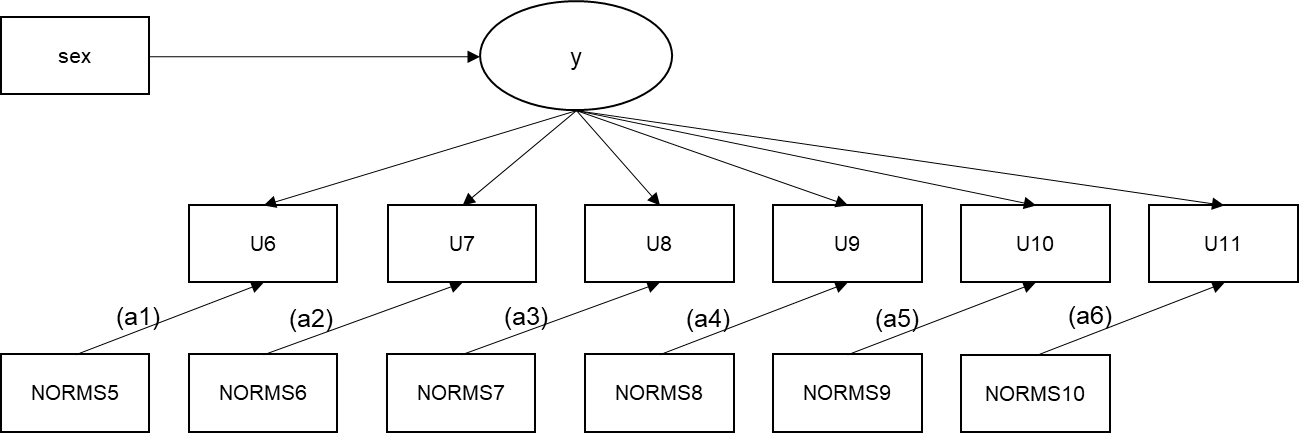
U8$1 3.409 0.195 17.439 0.000

U9$1 3.704 0.224 16.534 0.000

U10$1 3.683 0.228 16.167 0.000

U11$1 3.590 0.238 15.078 0.000

# INPUT EX 3b: Time-varying Effects



*The denotations (a1)-(a6) indicate that in this version of the model, the effect of NORMS is estimated independently at each lag.*

TITLE: EX 3b: TIME-VARYING PREDICTORS, TIME-VARYING (INDEPENDENT) EFFECT, CONTROLLING FOR TIME-INVARIANT COVARIATE

DATA:

FILE IS SurvivalExampleData4.dat;

VARIABLE:

NAMES ARE ID

MJ1 - MJ26

DRK1 - DRK26

puff1 - puff26

sex

hisp

white

cohort

T1MJ

POS1-POS17 !Positive Expectancies (continuous)

NEG1-NEG17 !Negative Expectancies (continuous)

AVAIL1-AVAIL17 !Alcohol Availability (Dichotomous)

NORMS1-NORMS17 !Perceived Norms (continuous)

;

USEOBSERVATIONS= (MJ2 NE 1) AND (MJ3 NE 1) AND (MJ4 NE 1) AND (MJ5 NE 1);

USEVARIABLES = sex NORMS5-NORMS10 U6-U11;

Missing = all(999);

CATEGORICAL = U6-U11;

DSURVIVAL = U6-U11;

DATA SURVIVAL:

NAMES = MJ6-MJ11;

CUTPOINT = 0;

BINARY = U6-U11;

ANALYSIS:

process=4;

coverage=0;

ESTIMATOR = MLR;

ALGORITHM = integration;

integration = Montecarlo;

MODEL:

y by U6-U11@1;

y on sex;

y@0;

U6 on NORMS5 (a1);

U7 on NORMS6 (a2);

U8 on NORMS7 (a3);

U9 on NORMS8 (a4);

U10 on NORMS9 (a5);

U11 on NORMS10 (a6);

!MODEL TEST requests a post-hoc Wald test for the difference in -2LL between the time-invariant model and time-varying model given the constraints imposed underneath. It does not affect model parameters, it simply adds a section in Results that provides information about whether freeing these values as a set significantly improves or decreases model fit.

MODEL TEST:

a1=a2; !All constraints must be listed separately with a single parameter on the left-hand side.

a2=a3;

a3=a4;

a4=a5;

a5=a6;

# OUTPUT EX 3b

EX 3b: TIME-VARYING PREDICTORS, TIME-VARYING EFFECT, CONTROLLING FOR TIME-INVAR

SUMMARY OF ANALYSIS

Number of groups 1

Number of observations 979

Number of dependent variables 6

Number of independent variables 7

Number of continuous latent variables 1

MODEL FIT INFORMATION

Number of Free Parameters 40

Loglikelihood

H0 Value -12290.357

H0 Scaling Correction Factor 1.4272

for MLR

Information Criteria

Akaike (AIC) 24660.713

Bayesian (BIC) 24856.174

Sample-Size Adjusted BIC 24729.134

(n\* = (n + 2) / 24)

Wald Test of Parameter Constraints

Value 4.277

Degrees of Freedom 5

P-Value 0.5103

!The Wald test of constraints requested by the MODEL TEST (matching model 2c) indicate that the PH !assumption is satisfied for the time-invariant (lagged) effect of the time-varying predictors.

!Note that this is only compared to completely independently estimated parameters.

MODEL RESULTS

Two-Tailed

Estimate S.E. Est./S.E. P-Value

Y BY

U6 1.000 0.000 999.000 999.000

U7 1.000 0.000 999.000 999.000

U8 1.000 0.000 999.000 999.000

U9 1.000 0.000 999.000 999.000

U10 1.000 0.000 999.000 999.000

U11 1.000 0.000 999.000 999.000

Y ON

SEX -0.311 0.154 -2.013 0.044

U6 ON

NORMS5 0.276 0.180 1.530 0.126

U7 ON

NORMS6 0.312 0.166 1.876 0.061

U8 ON

NORMS7 0.121 0.033 3.668 0.000

U9 ON

NORMS8 0.137 0.033 4.122 0.000

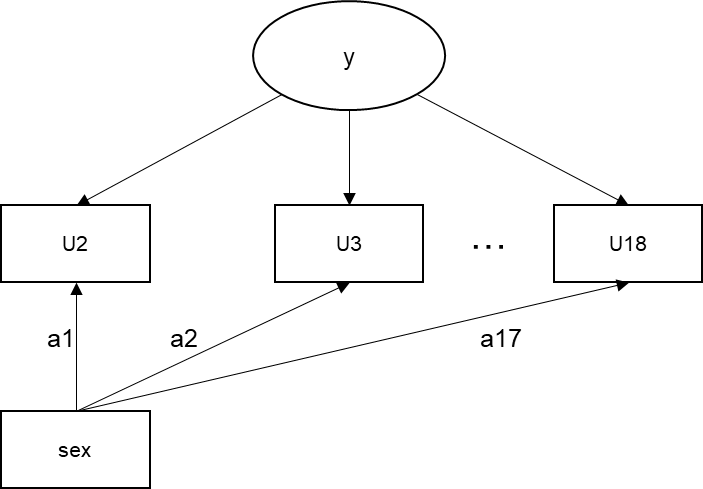
U10 ON

NORMS9 0.108 0.033 3.239 0.001

U11 ON

NORMS10 0.070 0.037 1.895 0.058

# INPUT EX 4a: Test of Proportional Hazards assumption - Independent



TITLE: EX 2: Testing the Proportional Hazards assumption - Independent Effects.

!In the independent (time-varying) effects model, the indicators are regressed separately on sex. Therefore, we allow for the possibility that sex affects probabilities at a certain time point (e.g. U2), but not on others (e.g. U3).

!This option might be best suited to situations where there are few timepoints for initiation data !(e.g. 5 or less).

!What makes this potentially misleading, is that the number of parameters increases rapidly, !meaning that any model fit test is highly likely to reject the more complex model.

!A parallel method for testing the Proportional Hazards assumption is to test for linear (and !quadratic) change in hazard over time. See Example 5b.

DATA:

FILE IS SurvivalExampleData.dat;

VARIABLE:

NAMES ARE ID

MJ1 - MJ26

DRK1 - DRK26

puff1 - puff26

sex

hisp

white

cohort;

USEVARIABLES = sex U1-U18;

Missing = all(999);

CATEGORICAL = U1-U18;

DSURVIVAL = U1-U18;

DATA SURVIVAL:

NAMES = MJ1-MJ18;

CUTPOINT = 0;

BINARY = U1-U18;

DEFINE:

sex = CENTER sex (GRANDMEAN);

ANALYSIS:

process=4;

coverage=0;

ESTIMATOR = MLR;

MODEL:

yi by U1-U18@1;

yi@0;

!u1 on sex; !Sparse data causes model to not identify unless excluded.

u2 on sex (a1);

u3 on sex (a2);

u4 on sex (a3);

u5 on sex (a4);

u6 on sex (a5);

u7 on sex (a6);

u8 on sex (a7);

u9 on sex (a8);

u10 on sex (a0);

u11 on sex (a10);

u12 on sex (a11);

u13 on sex (a12);

u14 on sex (a13);

u15 on sex (a14);

u16 on sex (a15);

u17 on sex (a16);

!Using the MODEL TEST command, we generate a Wald test indicating whether a single parameter estimating sex’s effect on the outcome constitutes a more effective, parsimonious model than the model here, where sex has an independent effect on each outcome.

MODEL TEST:

a1=a2;

a2=a3;

a3=a4;

a4=a5;

a5=a6;

a6=a7;

a7=a8;

a8=a9;

a9=a10;

a10=a11;

a11=a12;

a12=a13;

a13=a14;

a14=a15;

a15=a16;

!Another Mplus option is the MODEL CONSTRAINT option. Rather than testing the model against the constraints under MODEL TEST, MODEL CONSTRAINT simply imposes those constraints on the main model. In the syntax above, if MODEL TEST were replaced with MODEL CONSTRAINT, the estimate for sex would be identical to the one from MODEL 1a, but for the exclusion of u1 and u18 as outcomes.

# OUTPUT EX 4a

EX 4a: Testing the Proportional Hazards assumption - Independent Effects.

SUMMARY OF ANALYSIS

Number of groups 1

Number of observations 1023

Number of dependent variables 18

Number of independent variables 1

Number of continuous latent variables 1

MODEL FIT INFORMATION

Number of Free Parameters 35

!The steep increase in number of freely estimated parameters is unlikely to be justified by the increase in model fit.

Loglikelihood

H0 Value -1710.788

H0 Scaling Correction Factor 1.0000

for MLR

!The -2LL difference between this model and Model 1a, with a (35-19)=16 difference in parameters for the LRT, can be used to calculate whether the model fit is improved then the effect of sex is modeled independent of time.

Note that we use maximum likelihood with robust standard errors (MLR) as estimator using the Yuan-Bentler robust chi-square and sandwich standard errors (Yuan & Bentler, 1998). Note that MLR does not change parameter estimates, it only adjusts the standard errors of model parameters. In order to interpret loglikelihood differences using the loglikelihood difference test, calculate cd, TRd, and df according to the instructions on <https://stats.idre.ucla.edu/mplus/faq/how-can-i-compute-a-chi-square-test-for-nested-models-with-the-mlr-or-mlm-estimators/>, under ‘A test using the log-likelihood. The easiest way to use these numbers to calculate the *p*-value corresponding to the test, is to use the Excel function =chidist(df,TRd). It is always recommended to use MLR when available.

Information Criteria

Akaike (AIC) 3489.576

Bayesian (BIC) 3657.213

Sample-Size Adjusted BIC 3549.225

(n\* = (n + 2) / 24)

!The Wald test, a test similar to the loglikelihood difference test in function, automatically tests the model against the model with constraints imposed as described under MODEL TEST.

Wald Test of Parameter Constraints

Value 11.432

Degrees of Freedom 15

P-Value 0.7214

MODEL RESULTS

Two-Tailed

Estimate S.E. Est./S.E. P-Value

YI BY

U1 1.000 0.000 999.000 999.000

U2 1.000 0.000 999.000 999.000

U3 1.000 0.000 999.000 999.000

U4 1.000 0.000 999.000 999.000

U5 1.000 0.000 999.000 999.000

U6 1.000 0.000 999.000 999.000

U7 1.000 0.000 999.000 999.000

U8 1.000 0.000 999.000 999.000

U9 1.000 0.000 999.000 999.000

U10 1.000 0.000 999.000 999.000

U11 1.000 0.000 999.000 999.000

U12 1.000 0.000 999.000 999.000

U13 1.000 0.000 999.000 999.000

U14 1.000 0.000 999.000 999.000

U15 1.000 0.000 999.000 999.000

U16 1.000 0.000 999.000 999.000

U17 1.000 0.000 999.000 999.000

U18 1.000 0.000 999.000 999.000

U2 ON

SEX 0.089 1.002 0.088 0.930

U3 ON

SEX -0.249 0.589 -0.422 0.673

U4 ON

SEX 0.095 0.636 0.150 0.881

U5 ON

SEX 0.102 0.476 0.214 0.830

U6 ON

SEX 0.277 0.415 0.669 0.504

U7 ON

SEX -0.540 0.378 -1.426 0.154

U8 ON

SEX -0.372 0.336 -1.107 0.268

U9 ON

SEX -1.016 0.442 -2.297 0.022

U10 ON

SEX -0.137 0.381 -0.359 0.720

U11 ON

SEX -0.173 0.364 -0.474 0.636

U12 ON

SEX -0.639 0.349 -1.830 0.067

U13 ON

SEX -0.616 0.417 -1.477 0.140

U14 ON

SEX -0.264 0.374 -0.704 0.481

U15 ON

SEX 0.223 0.366 0.609 0.543

U16 ON

SEX -0.031 0.405 -0.077 0.938

U17 ON

SEX 0.203 0.435 0.466 0.641

Thresholds

U1$1 6.930 1.000 6.926 0.000

U2$1 5.582 0.708 7.879 0.000

U3$1 4.316 0.380 11.342 0.000

U4$1 4.641 0.449 10.327 0.000

U5$1 4.029 0.336 11.982 0.000

U6$1 3.792 0.305 12.439 0.000

U7$1 3.093 0.223 13.864 0.000

U8$1 2.886 0.210 13.758 0.000

U9$1 2.954 0.224 13.199 0.000

U10$1 3.165 0.255 12.401 0.000

U11$1 2.987 0.242 12.366 0.000

U12$1 2.513 0.204 12.326 0.000

U13$1 2.797 0.243 11.519 0.000

U14$1 2.697 0.244 11.075 0.000

U15$1 2.801 0.266 10.534 0.000

U16$1 2.777 0.275 10.082 0.000

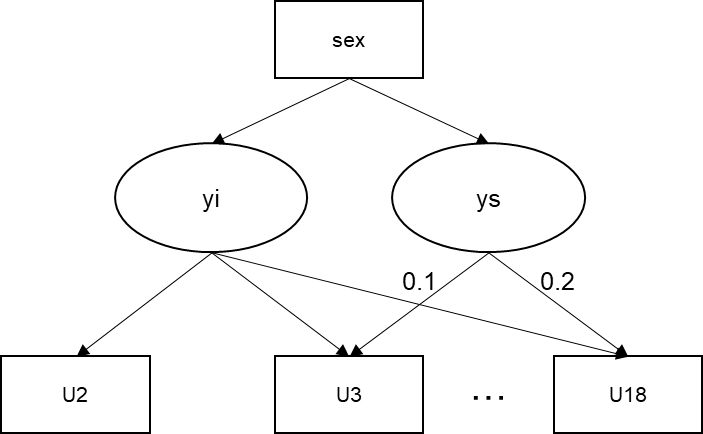
U17$1 2.559 0.300 8.539 0.000

U18$1 3.320 0.588 5.650 0.000

Residual Variances

YI 0.000 0.000 999.000 999.000

# INPUT EX 4b: Test of Proportional Hazards assumption - Linear



TITLE: EXAMPLE 4b: Testing the Proportional Hazards assumption.

DATA:

FILE IS SurvivalExampleData.dat;

VARIABLE:

NAMES ARE ID

MJ1 - MJ26

DRK1 - DRK26

puff1 - puff26

sex !Categorical variable with values 0 and 1

hisp !Categorical variable with values 0 and 1

white !Categorical variable with values 0 and 1

cohort !Categorical variable with values 1-5

;

USEVARIABLES = sex U1-U18;

Missing = all(999);

CATEGORICAL = U1-U18;

DSURVIVAL = U1-U18;

DATA SURVIVAL:

NAMES = MJ1-MJ18; !Source data - we use raw onset data on marijuana use.

CUTPOINT = 0; !Tells Mplus to discriminate between 0 and 1 for event

BINARY = U1-U18; !New transformed data for Model statement

ANALYSIS:

process=4;

coverage=0;

ESTIMATOR = MLR;

MODEL:

yi by U1-U18@1;

ys by U1@0 U2@.1 U3@.2 U4@.3 U5@.4 U6@.5 U7@.6 U8@.7 U9@.8 U10@.9 U11@1 U12@1.1

U13@1.2 U14@1.3 U15@1.4 U16@1.5 U17@1.6 U18@1.7;

yi ys on sex;

yi with ys@0;

yi@0;

ys@0;

# OUTPUT EX 4b:

EXAMPLE 4b: Testing the Proportional Hazards assumption.

SUMMARY OF ANALYSIS

Number of groups 1

Number of observations 1023

Number of dependent variables 18

Number of independent variables 1

Number of continuous latent variables 2

MODEL FIT INFORMATION

Number of Free Parameters 20

Loglikelihood

H0 Value -1716.958

H0 Scaling Correction Factor 1.0023

for MLR

Information Criteria

Akaike (AIC) 3473.916

Bayesian (BIC) 3572.525

Sample-Size Adjusted BIC 3509.003

(n\* = (n + 2) / 24)

MODEL RESULTS

Two-Tailed

Estimate S.E. Est./S.E. P-Value

YI BY

U1 1.000 0.000 999.000 999.000

U2 1.000 0.000 999.000 999.000

U3 1.000 0.000 999.000 999.000

U4 1.000 0.000 999.000 999.000

U5 1.000 0.000 999.000 999.000

U6 1.000 0.000 999.000 999.000

U7 1.000 0.000 999.000 999.000

U8 1.000 0.000 999.000 999.000

U9 1.000 0.000 999.000 999.000

U10 1.000 0.000 999.000 999.000

U11 1.000 0.000 999.000 999.000

U12 1.000 0.000 999.000 999.000

U13 1.000 0.000 999.000 999.000

U14 1.000 0.000 999.000 999.000

U15 1.000 0.000 999.000 999.000

U16 1.000 0.000 999.000 999.000

U17 1.000 0.000 999.000 999.000

U18 1.000 0.000 999.000 999.000

YS BY

U1 0.000 0.000 999.000 999.000

U2 0.100 0.000 999.000 999.000

U3 0.200 0.000 999.000 999.000

U4 0.300 0.000 999.000 999.000

U5 0.400 0.000 999.000 999.000

U6 0.500 0.000 999.000 999.000

U7 0.600 0.000 999.000 999.000

U8 0.700 0.000 999.000 999.000

U9 0.800 0.000 999.000 999.000

U10 0.900 0.000 999.000 999.000

U11 1.000 0.000 999.000 999.000

U12 1.100 0.000 999.000 999.000

U13 1.200 0.000 999.000 999.000

U14 1.300 0.000 999.000 999.000

U15 1.400 0.000 999.000 999.000

U16 1.500 0.000 999.000 999.000

U17 1.600 0.000 999.000 999.000

U18 1.700 0.000 999.000 999.000

YI ON

SEX -0.317 0.273 -1.161 0.246

YS ON

SEX 0.096 0.264 0.364 0.716

Thresholds

U1$1 6.791 0.989 6.863 0.000

U2$1 5.403 0.506 10.670 0.000

U3$1 4.296 0.304 14.148 0.000

U4$1 4.467 0.324 13.796 0.000

U5$1 3.857 0.244 15.783 0.000

U6$1 3.533 0.211 16.749 0.000

U7$1 3.199 0.191 16.777 0.000

U8$1 2.935 0.172 17.042 0.000

U9$1 3.213 0.201 16.003 0.000

U10$1 3.124 0.194 16.092 0.000

U11$1 2.967 0.186 15.947 0.000

U12$1 2.678 0.174 15.405 0.000

U13$1 2.955 0.206 14.334 0.000

U14$1 2.728 0.196 13.948 0.000

U15$1 2.605 0.192 13.585 0.000

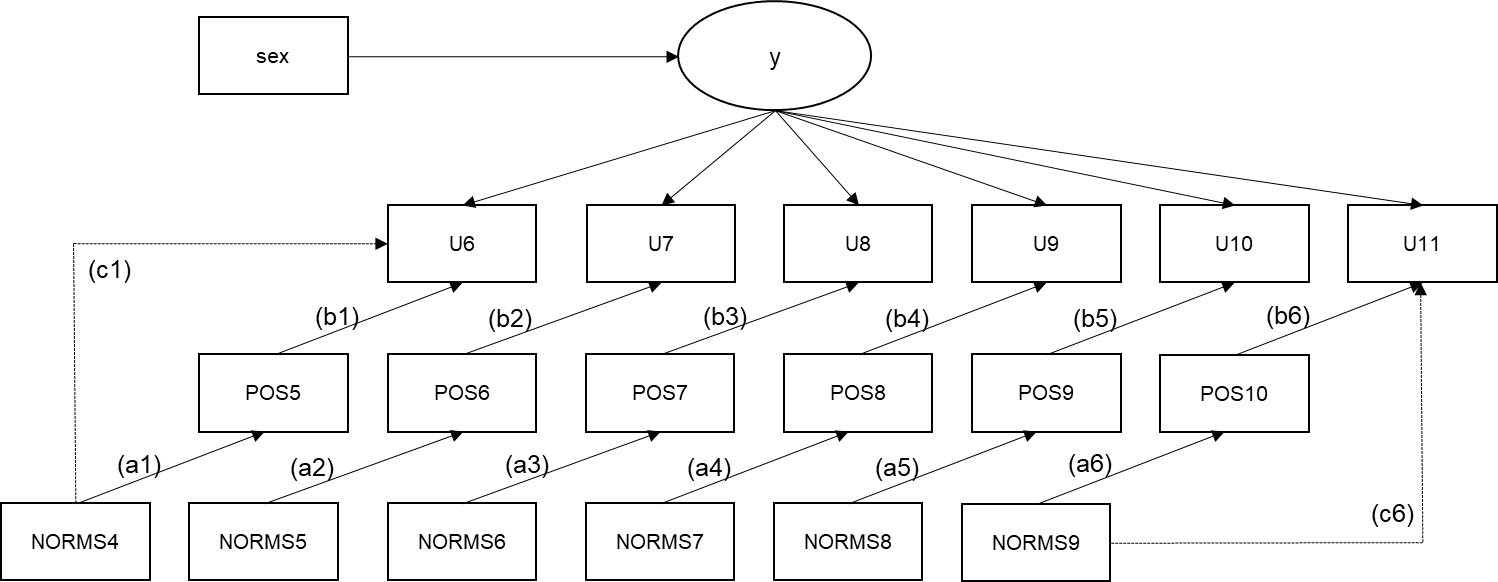
U16$1 2.714 0.215 12.634 0.000

U17$1 2.399 0.228 10.542 0.000

U18$1 3.255 0.589 5.526 0.000

# 

# INPUT EX 5a: Mediation



*Not shown: (c2-c5) direct paths from NORMS predictors to U outcomes.*

TITLE: EX 5a: TIME-VARYING MEDIATION IN SURVIVAL MODELING

!The following is an example of applying time-varying mediation in survival modeling, which is a complex strategy that requires a large amount of computation time. This is because significant mediation is typically established when the product of the indirect paths from the predictor to the mediator (the “a-path”) and from the mediator to the outcome (the “b-path”) is significant. However, the product of two estimates is ordinarily non-normal (see Preacher and Hayes, 2008). Therefore, we estimate this product 5000 times in bootstrapped samples drawn randomly with replacement from the data, to arrive at a non-normally distributed confidence interval of effect sizes.

DATA:

FILE IS SurvivalExampleData4.dat;

VARIABLE:

NAMES ARE ID

MJ1 - MJ26

DRK1 - DRK26

puff1 - puff26

sex !Categorical variable with values 0 and 1

hisp !Categorical variable with values 0 and 1

white !Categorical variable with values 0 and 1

cohort !Categorical variable with values 1-5

T1MJ !Variable indicating onset at time point T1

POS1-POS17 !Positive Alcohol Expectancies (continuous)

NEG1-NEG17 !Negative Expectancies (continuous)

AVAIL1-AVAIL17 !Alcohol Availability (Dichotomous)

NORMS1-NORMS17 !Perceived Norms (continuous)

;

USEOBSERVATIONS= (DRK2 NE 1) AND (DRK3 NE 1) AND (DRK4 NE 1) AND (DRK5 NE 1);!We add a time-varying mediator assessed at a lag-1, with the distal predictor now assessed at lag-2.

USEVARIABLES = sex **NORMS4-NORMS9** **POS5-POS10** U6-U11;

Missing = all(999);

CATEGORICAL = U6-U11;

DSURVIVAL = U6-U11;

DATA SURVIVAL:

NAMES = DRK6- DRK11;

CUTPOINT = 0;

BINARY = U6-U11;

ANALYSIS:

process=4;

coverage=0;

!Bootstrapping estimates is not possible with MLR (since both are separate techniques of dealing with non-normality), but we do have to specify the estimator = ML.

ESTIMATOR = ML;

ALGORITHM = integration;

integration = Montecarlo;

BOOTSTRAP = 5000;

MODEL:

y by U6-U11@1;

y on sex;

y@0;

!To estimate the a-path of the mediation, the mediator is regressed on the distal predictor.

POS5 on NORMS4 (a1);

POS6 on NORMS5 (a2);

POS7 on NORMS6 (a3);

POS8 on NORMS7 (a4);

POS9 on NORMS8 (a5);

POS10 on NORMS9 (a6);

!For the b- and c’-path, the mediator and distal predictor are both regressed on the time-specific outcome.

U6 on POS5 (b1)

NORMS4 (c1);

U7 on POS6 (b2)

NORMS5 (c2);

U8 on POS7 (b3)

NORMS6 (c3);

U9 on POS8 (b4)

NORMS7 (c4);

U10 on POS9 (b5)

NORMS8 (c5);

U11 on POS10 (b6)

NORMS9 (c6);

!Since the mediation is time-varying, it is estimated independently for each time point.

MODEL CONSTRAINT:

!MODEL CONSTRAINT allows us to specify additional parameters to be calculated based on post-hoc derivations of already estimated parameters (the model loglikelihood does not change due to their inclusion). One such parameter is the ‘bootstrapped joint indirect effect’, the product of the a-path and b-path. Bootstrapping is required because the joint distribution of the product of these parameters is almost by definition non-normal.

new(ind1-ind6);

ind1=a1\*b1;

ind2=a2\*b2;

ind3=a3\*b3;

ind4=a4\*b4;

ind5=a5\*b5;

ind6=a6\*b6;

output:

!To obtain the confidence interval of the bootstrapped joint indirect effect, we request Mplus to output the confidence interval based on bias-corrected bootstrapped effect ranges.

cinterval(bcbootstrap);

# OUTPUT EX 5a

EX 5a: TIME-VARYING MEDIATION IN SURVIVAL MODELING

SUMMARY OF ANALYSIS

Number of groups 1

Number of observations 979

Number of dependent variables 12

Number of independent variables 7

Number of continuous latent variables 1

!Estimator now is ML, bootstrap draws equal the amount requested, and the integration points are reported.

Estimator ML

Information matrix OBSERVED

Optimization Specifications for the Quasi-Newton Algorithm for

Continuous Outcomes

Maximum number of iterations 100

Convergence criterion 0.100D-05

Optimization Specifications for the EM Algorithm

Maximum number of iterations 500

Convergence criteria

Loglikelihood change 0.100D-02

Relative loglikelihood change 0.100D-05

Derivative 0.100D-02

Optimization Specifications for the M step of the EM Algorithm for

Categorical Latent variables

Number of M step iterations 1

M step convergence criterion 0.100D-02

Basis for M step termination ITERATION

Optimization Specifications for the M step of the EM Algorithm for

Censored, Binary or Ordered Categorical (Ordinal), Unordered

Categorical (Nominal) and Count Outcomes

Number of M step iterations 1

M step convergence criterion 0.100D-02

Basis for M step termination ITERATION

Maximum value for logit thresholds 15

Minimum value for logit thresholds -15

Minimum expected cell size for chi-square 0.100D-01

Maximum number of iterations for H1 2000

Convergence criterion for H1 0.100D-03

Number of bootstrap draws

Requested 5000

Completed 5000

Optimization algorithm EMA

Integration Specifications

Type MONTECARLO

Number of integration points 3000

Dimensions of numerical integration 12

Adaptive quadrature ON

Monte Carlo integration seed 0

Link LOGIT

Cholesky OFF

MODEL FIT INFORMATION

Number of Free Parameters 64

Loglikelihood

H0 Value -24246.442

Information Criteria

Akaike (AIC) 48620.884

Bayesian (BIC) 48934.660

Sample-Size Adjusted BIC 48731.393

MODEL RESULTS

Two-Tailed

Estimate S.E. Est./S.E. P-Value

Y BY

U6 1.000 0.000 999.000 999.000

U7 1.000 0.000 999.000 999.000

U8 1.000 0.000 999.000 999.000

U9 1.000 0.000 999.000 999.000

U10 1.000 0.000 999.000 999.000

U11 1.000 0.000 999.000 999.000

Y ON

SEX -0.541 0.134 -4.027 0.000

!Here start the estimates for a-paths (these are estimates with normally distributed standard errors, not the bootstrapped ones; those are reported further down).

POS5 ON

PNORM4 0.752 0.318 2.364 0.018

POS6 ON

PNORM5 0.170 0.239 0.710 0.478

POS7 ON

PNORM6 0.066 0.228 0.288 0.773

POS8 ON

PNORM7 0.116 0.104 1.118 0.264

POS9 ON

PNORM8 0.134 0.085 1.562 0.118

POS10 ON

PNORM9 0.257 0.085 3.011 0.003!

Here start the b- and the c’-paths. The lack of significance of the effects of POS as a predictor suggest that POS is not related to subsequent initiation after controlling for NORMS (x).

U6 ON

POS5 0.020 0.034 0.585 0.559

PNORM4 0.434 0.155 2.797 0.005

U7 ON

POS6 0.021 0.047 0.440 0.660

PNORM5 0.437 0.154 2.834 0.005

U8 ON

POS7 0.034 0.035 0.971 0.332

PNORM6 0.702 0.275 2.551 0.011

U9 ON

POS8 0.029 0.025 1.175 0.240

PNORM7 0.111 0.034 3.251 0.001

U10 ON

POS9 0.054 0.024 2.289 0.022

PNORM8 0.075 0.040 1.870 0.062

U11 ON

POS10 0.097 0.056 1.753 0.080

PNORM9 0.017 0.036 0.474 0.635

Thresholds

U6$1 4.159 0.841 4.948 0.000

U7$1 4.130 1.217 3.392 0.001

U8$1 5.031 1.102 4.565 0.000

U9$1 3.825 0.634 6.036 0.000

U10$1 4.092 0.638 6.412 0.000

U11$1 4.500 1.579 2.851 0.004

New/Additional Parameters

IND1 0.015 0.027 0.556 0.578

IND2 0.004 0.016 0.225 0.822

IND3 0.002 0.010 0.212 0.832

IND4 0.003 0.005 0.659 0.510

IND5 0.007 0.006 1.237 0.216

IND6 0.025 0.016 1.527 0.127

!This is where the bootstrapped confidence intervals start. For intervals matching an alpha-level of .05, we are looking for the lower 2.5% to the higher 2.5% bound NOT to contain zero.

CONFIDENCE INTERVALS OF MODEL RESULTS

Lower .5% Lower 2.5% Lower 5% Estimate Upper 5% Upper 2.5% Upper .5%

Y ON

SEX -0.925 -0.808 -0.757 -0.541 -0.328 -0.289 -0.221

POS5 ON

PNORM4 -0.018 0.216 0.303 0.752 1.345 1.417 1.593

POS6 ON

PNORM5 -0.462 -0.382 -0.294 0.170 0.540 0.618 0.736

POS7 ON

PNORM6 -0.532 -0.397 -0.326 0.066 0.441 0.508 0.639

POS8 ON

PNORM7 -0.159 -0.097 -0.056 0.116 0.288 0.317 0.395

POS9 ON

PNORM8 -0.061 -0.008 0.023 0.134 0.298 0.343 0.374

POS10 ON

PNORM9 0.046 0.083 0.105 0.257 0.399 0.428 0.469

U6 ON

POS5 -0.088 -0.047 -0.039 0.020 0.071 0.083 0.109

PNORM4 -0.022 0.112 0.158 0.434 0.696 0.748 0.821

U7 ON

POS6 -0.105 -0.066 -0.052 0.021 0.092 0.115 0.178

PNORM5 0.055 0.154 0.203 0.437 0.702 0.756 0.928

U8 ON

POS7 -0.045 -0.025 -0.016 0.034 0.100 0.119 0.141

PNORM6 0.201 0.290 0.354 0.702 1.090 1.212 1.442

U9 ON

POS8 -0.030 -0.020 -0.012 0.029 0.069 0.077 0.091

PNORM7 0.012 0.045 0.054 0.111 0.165 0.179 0.193

U10 ON

POS9 -0.007 0.011 0.017 0.054 0.089 0.098 0.122

PNORM8 -0.041 -0.014 0.003 0.075 0.136 0.146 0.171

U11 ON

POS10 -0.007 0.020 0.028 0.097 0.183 0.208 0.260

PNORM9 -0.091 -0.051 -0.037 0.017 0.075 0.080 0.105

!Here are the parameters of greatest interest: The bootstrapped joint estimates. The fact that the range of estimates (lower 95% to upper 95%) for ‘IND5’ and ‘IND6’ do not contain zero, suggests that significant mediation occurs at these time points.

New/Additional Parameters

IND1 -0.049 -0.026 -0.016 0.015 0.067 0.081 0.110

IND2 -0.033 -0.016 -0.010 0.004 0.038 0.049 0.080

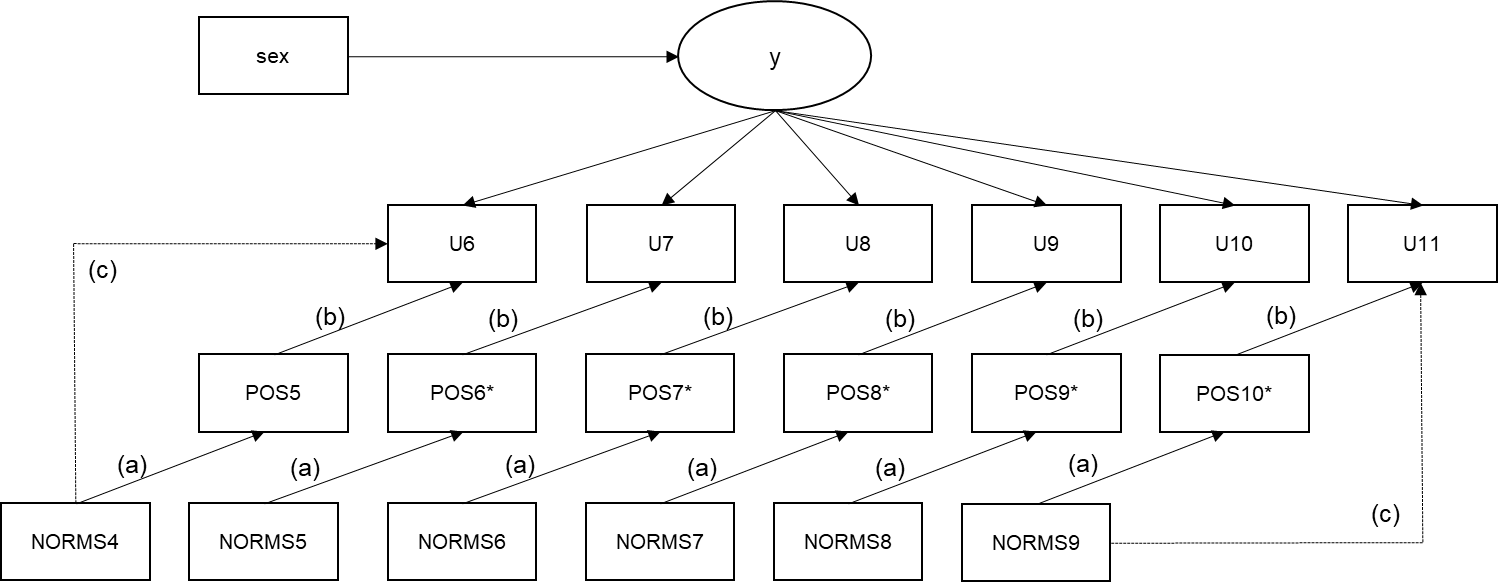
IND3 -0.020 -0.011 -0.007 0.002 0.030 0.039 0.054

IND4 -0.007 -0.003 -0.002 0.003 0.016 0.018 0.023

IND5 -0.004 0.000 0.001 0.007 0.021 0.031 0.046

IND6 0.000 0.006 0.008 0.025 0.054 0.061 0.089

# INPUT EX 5b: Time-invariant Mediation



*Pathways at different time points are constrained to have an equal effect. \*Observations on variables with an asterisk are removed if initiation has occurred concurrently or prior to that observation.*

TITLE: EX 5b: TIME-INVARIANT EFFECT OF TIME-VARYING MEDIATOR IN SURVIVAL MODELING

DATA:

FILE IS SurvivalExampleData4.dat;

VARIABLE:

NAMES ARE ID

MJ1 - MJ26

DRK1 - DRK26

puff1 - puff26

sex !Categorical variable with values 0 and 1

hisp !Categorical variable with values 0 and 1

white !Categorical variable with values 0 and 1

cohort !Categorical variable with values 1-5

T1MJ !Variable indicating onset at time point T1

POS1-POS17 !Positive Alcohol Expectancies (continuous)

NEG1-NEG17 !Negative Expectancies (continuous)

AVAIL1-AVAIL17 !Alcohol Availability (Dichotomous)

NORMS1-NORMS17 !Perceived Norms (continuous)

;

USEOBSERVATIONS= (DRK2 NE 1) AND (DRK3 NE 1) AND (DRK4 NE 1) AND (DRK5 NE 1);

USEVARIABLES = sex NORMS4-NORMS9 POS5-POS10 U6-U11;

Missing = all(999);

CATEGORICAL = U6-U11;

DSURVIVAL = U6-U11;

DATA SURVIVAL:

NAMES = DRK6-DRK11;

CUTPOINT = 0;

BINARY = U6-U11;

ANALYSIS:

process=4;

coverage=0;

ESTIMATOR = ML;

ALGORITHM = integration;

integration = Montecarlo (250);

BOOTSTRAP = 2000;

!When defining a time-invariant effect, there is the potential for violation of temporal order because the a-path is estimated equally at each time point, using all available data on the mediator POS. This is an issue, because POS is also assessed when onset has already occurred. To correct for this issue, we eliminate all observations for POS that are performed when onset has already occurred.

DEFINE:

if DRK6==1 then POS6=\_MISSING;

if DRK7==1 OR DRK6==1 then POS7=\_MISSING;

if DRK8==1 OR DRK7==1 OR DRK6==1 then POS8=\_MISSING;

if DRK9==1 OR DRK8==1 OR DRK7==1 OR DRK6==1 then POS9=\_MISSING;

if DRK10==1 OR DRK9==1 OR DRK8==1 OR DRK7==1 OR DRK6==1 then POS10=\_MISSING;

MODEL:

y by U6-U11@1;

y on sex;

y@0;

!By naming parameters the same, we constrain them to be equal.

POS5 on NORMS4 (a);

POS6 on NORMS5 (a);

POS7 on NORMS6 (a);

POS8 on NORMS7 (a);

POS9 on NORMS8 (a);

POS10 on NORMS9 (a);

U6 on POS5 (b)

NORMS4 (c);

U7 on POS6 (b)

NORMS5 (c);

U8 on POS7 (b)

NORMS6 (c);

U9 on POS8 (b)

NORMS7 (c);

U10 on POS9 (b)

NORMS8 (c);

U11 on POS10 (b)

NORMS9 (c);

!A single bootstrapped indirect effect is estimated, based on time-invariant effects a and b, controlling for time-invariant effect c’ (and time-invariant covariate sex).

MODEL CONSTRAINT:

new(ind1);

ind1=a\*b;

output:

cinterval(bcbootstrap);

# OUTPUT EX 5b

EX 5b: TIME-INVARIANT EFFECT OF TIME-VARYING MEDIATOR IN SURVIVAL MODELING

MODEL FIT INFORMATION

Number of Free Parameters 49

Loglikelihood

H0 Value -23685.465

Information Criteria

Akaike (AIC) 47468.931

Bayesian (BIC) 47709.165

Sample-Size Adjusted BIC 47553.539

(n\* = (n + 2) / 24)

MODEL RESULTS

Two-Tailed

Estimate S.E. Est./S.E. P-Value

Y BY

U6 1.000 0.000 999.000 999.000

U7 1.000 0.000 999.000 999.000

U8 1.000 0.000 999.000 999.000

U9 1.000 0.000 999.000 999.000

U10 1.000 0.000 999.000 999.000

U11 1.000 0.000 999.000 999.000

Y ON

SEX -0.552 0.133 -4.156 0.000

POS5 ON

PNORM4 0.187 0.062 3.019 0.003

POS6 ON

PNORM5 0.187 0.062 3.019 0.003

POS7 ON

PNORM6 0.187 0.062 3.019 0.003

POS8 ON

PNORM7 0.187 0.062 3.019 0.003

POS9 ON

PNORM8 0.187 0.062 3.019 0.003

POS10 ON

PNORM9 0.187 0.062 3.019 0.003

U6 ON

POS5 0.045 0.013 3.483 0.000

PNORM4 0.086 0.018 4.687 0.000

U7 ON

POS6 0.045 0.013 3.483 0.000

PNORM5 0.086 0.018 4.687 0.000

U8 ON

POS7 0.045 0.013 3.483 0.000

PNORM6 0.086 0.018 4.687 0.000

U9 ON

POS8 0.045 0.013 3.483 0.000

PNORM7 0.086 0.018 4.687 0.000

U10 ON

POS9 0.045 0.013 3.483 0.000

PNORM8 0.086 0.018 4.687 0.000

U11 ON

POS10 0.045 0.013 3.483 0.000

PNORM9 0.086 0.018 4.687 0.000

Thresholds

U6$1 4.292 0.366 11.724 0.000

U7$1 4.175 0.367 11.385 0.000

U8$1 4.038 0.360 11.205 0.000

U9$1 4.073 0.350 11.627 0.000

U10$1 3.883 0.353 11.016 0.000

U11$1 3.440 0.360 9.547 0.000

!The parameter for the indirect effect, with normally distributed standard errors, is significant (p=.019). However, as mentioned, the indirect effect does not have normally distributed standard errors.

New/Additional Parameters

IND1 0.008 0.004 2.343 0.019

CONFIDENCE INTERVALS OF MODEL RESULTS

Lower .5% Lower 2.5% Lower 5% Estimate Upper 5% Upper 2.5% Upper .5%

Y BY

U6 1.000 1.000 1.000 1.000 1.000 1.000 1.000

U7 1.000 1.000 1.000 1.000 1.000 1.000 1.000

U8 1.000 1.000 1.000 1.000 1.000 1.000 1.000

U9 1.000 1.000 1.000 1.000 1.000 1.000 1.000

U10 1.000 1.000 1.000 1.000 1.000 1.000 1.000

U11 1.000 1.000 1.000 1.000 1.000 1.000 1.000

Y ON

SEX -0.936 -0.832 -0.782 -0.552 -0.350 -0.311 -0.236

POS5 ON

PNORM4 0.052 0.090 0.108 0.187 0.315 0.332 0.380

POS6 ON

PNORM5 0.052 0.090 0.108 0.187 0.315 0.332 0.380

POS7 ON

PNORM6 0.052 0.090 0.108 0.187 0.315 0.332 0.380

POS8 ON

PNORM7 0.052 0.090 0.108 0.187 0.315 0.332 0.380

POS9 ON

PNORM8 0.052 0.090 0.108 0.187 0.315 0.332 0.380

POS10 ON

PNORM9 0.052 0.090 0.108 0.187 0.315 0.332 0.380

U6 ON

POS5 0.016 0.023 0.027 0.045 0.070 0.075 0.080

PNORM4 0.027 0.044 0.054 0.086 0.115 0.119 0.127

U7 ON

POS6 0.016 0.023 0.027 0.045 0.070 0.075 0.080

PNORM5 0.027 0.044 0.054 0.086 0.115 0.119 0.127

U8 ON

POS7 0.016 0.023 0.027 0.045 0.070 0.075 0.080

PNORM6 0.027 0.044 0.054 0.086 0.115 0.119 0.127

U9 ON

POS8 0.016 0.023 0.027 0.045 0.070 0.075 0.080

PNORM7 0.027 0.044 0.054 0.086 0.115 0.119 0.127

U10 ON

POS9 0.016 0.023 0.027 0.045 0.070 0.075 0.080

PNORM8 0.027 0.044 0.054 0.086 0.115 0.119 0.127

U11 ON

POS10 0.016 0.023 0.027 0.045 0.070 0.075 0.080

PNORM9 0.027 0.044 0.054 0.086 0.115 0.119 0.127

Thresholds

U6$1 4.292 4.292 4.292 4.292 4.292 4.292 4.292

U7$1 4.175 4.175 4.175 4.175 4.175 4.175 4.175

U8$1 4.038 4.038 4.038 4.038 4.038 4.038 4.038

U9$1 4.073 4.073 4.073 4.073 4.073 4.073 4.073

U10$1 3.883 3.883 3.883 3.883 3.883 3.883 3.883

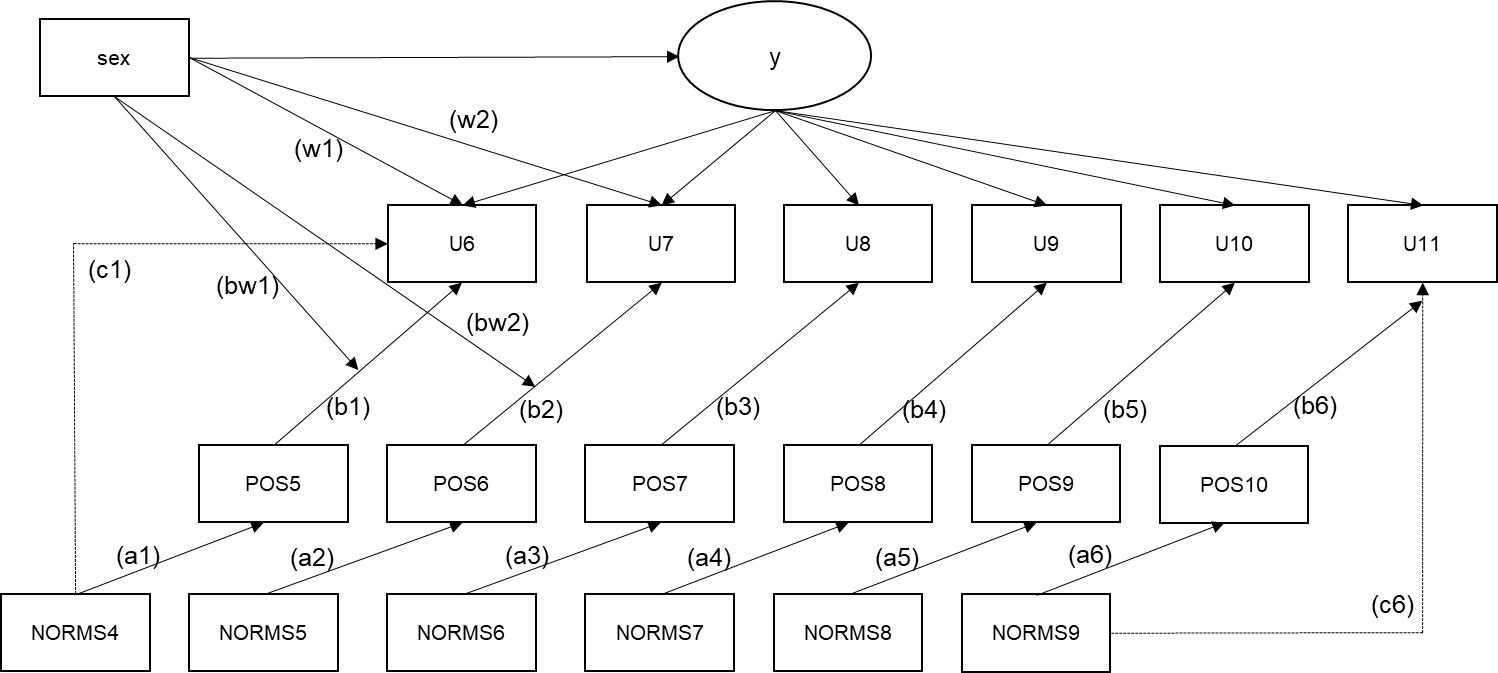
U11$1 3.440 3.440 3.440 3.440 3.440 3.440 3.440

!The estimated bootstrapped confidence interval for the mediated effect does not contain zero (the 99% bootstrapped CI is .002 to .028), therefore, we conclude that significant mediation occurs, when we examine mediation in a time-invariant way.

New/Additional Parameters

IND1 0.002 0.003 0.004 0.008 0.017 0.019 0.028

# INPUT EX 5c: Moderated Mediation



*Not shown: sex moderating pathways b3, b4, b5, b6, as well as bw3, bw4, bw5, and bw6. Moderation is shown here as arrows intersecting with other arrows. In actual syntax, for example, ‘bw1’ is a pathway reflecting the prediction of u6 by a new variable that is calculated as the product of sex and pos5.*

TITLE: EX 5c: MODERATED MEDIATION IN SURVIVAL MODELING

DATA:

FILE IS SurvivalExampleData4.dat;

VARIABLE:

NAMES ARE ID

MJ1 - MJ26

DRK1 - DRK26

puff1 - puff26

sex

hisp

white

cohort

T1MJ

POS1-POS17 !Positive Alcohol Expectancies (continuous)

NEG1-NEG17 !Negative Expectancies (continuous)

AVAIL1-AVAIL17 !Alcohol Availability (Dichotomous)

NORMS1-NORMS17 !Perceived Norms (continuous)

;

USEOBSERVATIONS= (DRK2 NE 1) AND (DRK3 NE 1) AND (DRK4 NE 1) AND (DRK5 NE 1);

USEVARIABLES =

sex

NORMS4-NORMS9

POS5-POS10

U6-U11

sexPos5-sexPos10;

Missing = all(999);

CATEGORICAL = U6-U11;

DSURVIVAL = U6-U11;

DATA SURVIVAL:

NAMES = DRK6-DRK11;

CUTPOINT = 0;

BINARY = U6-U11;

ANALYSIS:

process=4;

coverage=0;

ESTIMATOR = ML;

ALGORITHM = integration;

integration = Montecarlo;

BOOTSTRAP = 5000;

DEFINE:

STANDARDIZE POS5-POS10;

sexPos5=sex\*Pos5;

sexPos6=sex\*Pos6;

sexPos7=sex\*Pos7;

sexPos8=sex\*Pos8;

sexPos9=sex\*Pos9;

sexPos10=sex\*Pos10;

MODEL:

! y by U6-U11@1;

! y on sex;

! y@0;

POS5 on NORMS4 (a1);

POS6 on NORMS5 (a2);

POS7 on NORMS6 (a3);

POS8 on NORMS7 (a4);

POS9 on NORMS8 (a5);

POS10 on NORMS9 (a6);

U6 on POS5 (b1)

NORMS4 (c1)

sex (w1)

sexPos5 (bw1);

U7 on POS6 (b2)

NORMS5 (c2)

sex (w2)

sexPos6 (bw2);

U8 on POS7 (b3)

NORMS6 (c3)

sex (w3)

sexPos7 (bw3);

U9 on POS8 (b4)

NORMS7 (c4)

sex (w4)

sexPos8 (bw4);

U10 on POS9 (b5)

NORMS8 (c5)

sex (w5)

sexPos9 (bw5);

U11 on POS10 (b6)

NORMS9 (c6)

sex (w6)

sexPos10 (bw6);

MODEL CONSTRAINT:

new(ind10 ind11 ind20 ind21 ind30 ind31 ind40 ind41 ind50 ind51 ind60 ind61);

ind10=a1\*b1;

ind11=a1\*(b1+bw1);

\*The effect of dichotomous moderator ‘sex’ is calculated as applied to the b-path if sex =1. Hence, ind10 reflects the indirect effect 1 if sex is at value 0. Ind11 reflects the indirect effect 1 if sex is at value 1.

\*For continuous moderators, the most useful approach may be to calculate three indirect effects for three values of the moderator: -1 SD, mean, and +1SD. Predictors and moderators should be standardized if both are continuous.

ind20=a2\*b2;

ind21=a2\*(b2+bw2);

ind30=a3\*b3;

ind31=a3\*(b3+bw3);

ind40=a4\*b4;

ind41=a4\*(b4+bw4);

ind50=a5\*b5;

ind51=a5\*(b5+bw5);

ind60=a6\*b6;

ind61=a6\*(b6+bw6);

output:

cinterval(bcbootstrap);

# OUTPUT EX 5c

EX 5c: MODERATED MEDIATION IN SURVIVAL MODELING

SU SUMMARY OF ANALYSIS

Number of groups 1

Number of observations 995

Number of dependent variables 12

Number of independent variables 13

Number of continuous latent variables 0

MODEL FIT INFORMATION

Number of Free Parameters 138

Loglikelihood

H0 Value -20236.256

Information Criteria

Akaike (AIC) 40748.512

Bayesian (BIC) 41425.091

Sample-Size Adjusted BIC 40986.796

(n\* = (n + 2) / 24)

MODEL RESULTS

Two-Tailed

Estimate S.E. Est./S.E. P-Value

!The effect of the distal predictor is mostly positive, but only once significant.

POS5 ON

PNORM4 0.095 0.049 1.935 0.053

POS6 ON

PNORM5 0.028 0.035 0.816 0.415

POS7 ON

PNORM6 -0.003 0.032 -0.092 0.927

POS8 ON

PNORM7 0.016 0.013 1.166 0.243

POS9 ON

PNORM8 0.013 0.011 1.123 0.261

POS10 ON

PNORM9 0.034 0.012 2.835 0.005

!Time-specific regression now includes the effect of the mediator, the distal predictor, the moderator, and the interaction between the moderator and the mediator.

U6 ON

POS5 0.011 0.323 0.035 0.972

PNORM4 0.432 0.158 2.734 0.006

SEX -0.346 0.418 -0.828 0.408

SEXPOS5 0.452 0.475 0.952 0.341

U7 ON

POS6 0.621 0.701 0.886 0.376

PNORM5 0.405 0.157 2.574 0.010

SEX -0.332 0.593 -0.560 0.576

SEXPOS6 -0.522 1.082 -0.483 0.629

U8 ON

POS7 -0.015 0.634 -0.024 0.981

PNORM6 0.715 0.263 2.717 0.007

SEX -0.789 0.754 -1.046 0.296

SEXPOS7 0.231 1.330 0.173 0.862

U9 ON

POS8 0.128 0.291 0.440 0.660

PNORM7 0.113 0.035 3.188 0.001

SEX -0.733 0.376 -1.948 0.051

SEXPOS8 0.293 0.380 0.772 0.440

U10 ON

POS9 0.522 0.291 1.790 0.073

PNORM8 0.078 0.041 1.874 0.061

SEX -0.330 0.343 -0.963 0.336

SEXPOS9 -0.146 0.354 -0.411 0.681

U11 ON

POS10 0.912 0.726 1.256 0.209

PNORM9 0.014 0.043 0.329 0.742

SEX -0.578 0.372 -1.553 0.120

SEXPOS10 -0.432 0.836 -0.517 0.605

New/Additional Parameters

IND10 0.001 0.035 0.031 0.976

IND11 0.044 0.043 1.021 0.307

IND20 0.018 0.031 0.564 0.573

IND21 0.003 0.028 0.101 0.919

IND30 0.000 0.024 0.002 0.999

IND31 -0.001 0.027 -0.023 0.981

IND40 0.002 0.006 0.324 0.746

IND41 0.007 0.007 0.887 0.375

IND50 0.007 0.008 0.840 0.401

IND51 0.005 0.005 0.894 0.371

IND60 0.031 0.029 1.059 0.290

IND61 0.016 0.016 1.035 0.301

CONFIDENCE INTERVALS OF MODEL RESULTS

Lower .5% Lower 2.5% Lower 5% Estimate Upper 5% Upper 2.5% Upper .5%

POS5 ON

PNORM4 -0.032 0.000 0.016 0.095 0.179 0.194 0.220

POS6 ON

PNORM5 -0.062 -0.040 -0.032 0.028 0.086 0.094 0.114

POS7 ON

PNORM6 -0.099 -0.081 -0.067 -0.003 0.040 0.050 0.068

POS8 ON

PNORM7 -0.019 -0.011 -0.006 0.016 0.038 0.042 0.047

POS9 ON

PNORM8 -0.017 -0.010 -0.007 0.013 0.031 0.035 0.042

POS10 ON

PNORM9 0.002 0.008 0.011 0.034 0.052 0.056 0.064

U6 ON

POS5 -1.228 -0.805 -0.665 0.011 0.426 0.515 0.757

PNORM4 0.000 0.114 0.155 0.432 0.680 0.738 0.841

SEX -1.775 -1.225 -1.086 -0.346 0.305 0.397 0.633

SEXPOS5 -0.593 -0.319 -0.184 0.452 1.423 1.660 1.969

U7 ON

POS6 -0.740 -0.444 -0.254 0.621 2.188 2.433 2.718

PNORM5 -0.122 0.074 0.120 0.405 0.631 0.679 0.799

SEX -1.998 -1.245 -1.063 -0.332 0.599 0.740 1.369

SEXPOS6 -4.220 -2.862 -2.561 -0.522 0.894 1.179 1.766

U8 ON

POS7 -1.640 -1.305 -1.187 -0.015 0.819 0.985 1.316

PNORM6 0.194 0.306 0.367 0.715 1.103 1.207 1.458

SEX -4.077 -2.544 -2.103 -0.789 -0.069 0.077 0.307

SEXPOS7 -2.449 -1.696 -1.414 0.231 3.038 3.665 5.095

U9 ON

POS8 -0.655 -0.423 -0.320 0.128 0.627 0.723 0.913

PNORM7 0.009 0.039 0.053 0.113 0.167 0.178 0.195

SEX -1.850 -1.481 -1.334 -0.733 -0.130 -0.029 0.189

SEXPOS8 -0.598 -0.392 -0.277 0.293 1.006 1.137 1.460

U10 ON

POS9 -0.461 -0.208 -0.053 0.522 0.935 1.022 1.194

PNORM8 -0.043 -0.009 0.007 0.078 0.142 0.150 0.176

SEX -1.263 -0.976 -0.875 -0.330 0.254 0.371 0.538

SEXPOS9 -0.891 -0.663 -0.569 -0.146 0.719 0.931 1.184

U11 ON

POS10 -1.204 -1.158 -1.096 0.912 1.467 1.605 1.925

PNORM9 -0.117 -0.070 -0.054 0.014 0.085 0.099 0.124

SEX -1.675 -1.318 -1.184 -0.578 0.011 0.163 0.486

SEXPOS10 -1.663 -1.306 -1.125 -0.432 2.077 2.310 2.623

!Indirect effect parameters with a ‘0’ at the end, denote the effect when moderator = 0 (here: participant is female). Those with ‘1’ at the end denote the effect when moderator = 1 (and thus, the moderating effect on the b-path is applied). We infer that moderated mediation does not occur.

New/Additional Parameters

IND10 -0.136 -0.093 -0.074 0.001 0.044 0.056 0.089

IND11 -0.024 -0.005 0.002 0.044 0.178 0.208 0.270

IND20 -0.028 -0.011 -0.004 0.018 0.141 0.154 0.161

IND21 -0.065 -0.036 -0.026 0.003 0.061 0.082 0.139

IND30 -0.099 -0.053 -0.040 0.000 0.037 0.049 0.083

IND31 -0.119 -0.068 -0.047 -0.001 0.032 0.044 0.085

IND40 -0.011 -0.005 -0.003 0.002 0.018 0.022 0.031

IND41 -0.006 -0.002 0.000 0.007 0.027 0.032 0.051

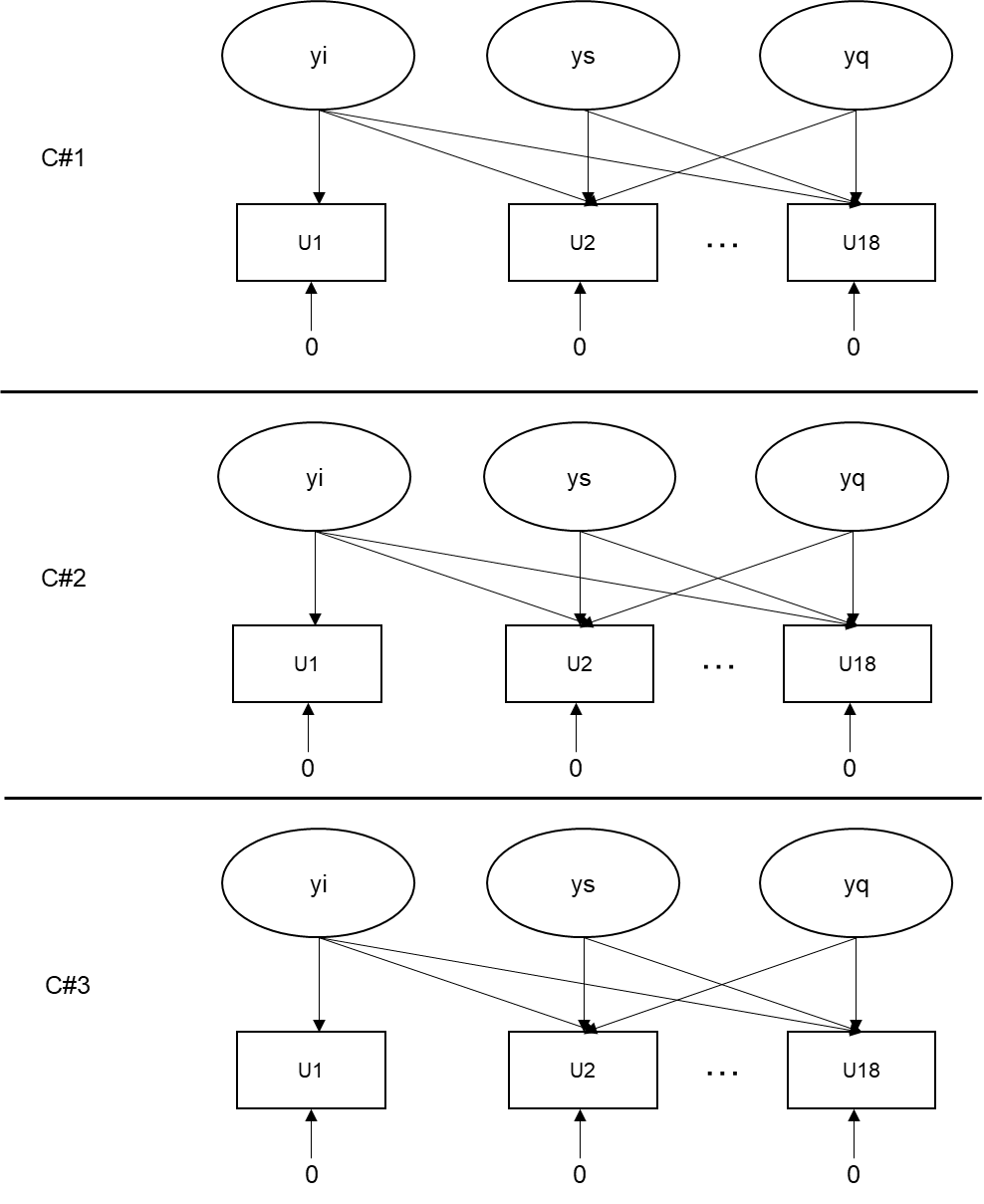
IND50 -0.014 -0.007 -0.004 0.007 0.022 0.025 0.034

IND51 -0.004 -0.001 0.000 0.005 0.021 0.024 0.032

IND60 -0.057 -0.043 -0.035 0.031 0.063 0.071 0.089

IND61 -0.029 -0.011 -0.005 0.016 0.040 0.047 0.067

# INPUT EX 6a: Multiple Survival Curves



TITLE: EX 6a: Multiple Event Process Survival Mixture Model

!This example is based on the syntax provided with Dean, Cole, & Bauer (2014). Briefly, this approach may be applied when initiation on multiple events is considered. Rather than estimate the discrete hazard at each time point, MEPSUM creates classes based on differences in the hazard function between classes, based on observed profiles of initiation within the sample.

!For instance, the approach below characterizes classes within our sample on initiation of marijuana use, full drink, cigarette puffing, and other drug use.

DATA:

FILE IS 20180416\_MEPSUM.dat;

VARIABLE:

NAMES ARE ID

MJ1-MJ26

DRK1-DRK26

PUFF1-PUFF26

OD1-OD26

sex hisp white cohort

;

USEVARIABLES =

MJ1 - MJ18

DRK1 - DRK18

OD2 OD5-OD17

puff2 - puff18

;

!We use the same initiation variables at before, but these have already been recoded to be structured in the way that DATA SURVIVAL would apply.

!Certain initiation indicators have been omitted here, as our descriptive data analyses (not shown here) had shown that no initiation took place during those time points.

IDvariable = ID;

Missing=all(999);

!the CLASSES command instructs Mplus to estimate a series of classes, C, here numbering 3. The differentiating features of classes C are listed below in the ‘Model:’ statement.

CLASSES = C(3);

CATEGORICAL =

MJ1 - MJ18

DRK1 - DRK18

puff2 - puff18

OD2 OD5-OD17;

Analysis:

TYPE = Mixture;

!MEPSUM solutions are relatively complex with a high proportion of local minima in the maximum likelihood function. Hence, the number of starts and final model solutions presented should be vastly increased (relative to the Mplus default) in order to ensure the true global minimum was identified. This is not a fool proof solution; in our models, a true global minimum was never established for the 4-class solution even when 1000 starts were used. The models identified, but the loglikelihood was not replicated.

STARTS = 500 100;

STITERATIONS = 25;

process=4;

coverage=0;

!OPTSEED=830392;

model:

%OVERALL% !Statements in the %OVERALL% section apply to all classes.

!Unlike in traditional survival modeling, the survival curve here is estimated within each class, in such a way that class-specific hazard is a function of an initial (intercept), linear (slope), and quadratic change over time. The hazard is estimated purely as a function; the residual thresholds for indicators are constrained to be equal to zero.

i1 s1 q1 | MJ1@.0 MJ2@.1 MJ3@.2 MJ4@.3 MJ5@.4 MJ6@.5 MJ7@.6 MJ8@.7 MJ9@.8

MJ10@.9 MJ11@1.0 MJ12@1.1 MJ13@1.2 MJ14@1.3 MJ15@1.4 MJ16@1.5

MJ17@1.6 MJ18@1.7;

!Note that the weights are kept on a small scale. The model becomes computationally more intensive if weights are larger. The relative scale of these weights are all that matters; hence, they can all be divided by 10, as was applied here.

i2 s2 q2 | Puff2@.1 Puff3@.2 Puff4@.3 Puff5@.4 Puff6@.5 Puff7@.6 Puff8@.7

Puff9@.8 Puff10@.9 Puff11@1.0 Puff12@1.1 Puff13@1.2 Puff14@1.3 Puff15@1.4

Puff16@1.5 Puff17@1.6 Puff18@1.7;

i3 s3 q3 | Drk1@.0 Drk2@.1 Drk3@.2 Drk4@.3 Drk5@.4 Drk6@.5 Drk7@.6 Drk8@.7 Drk9@.8

Drk10@.9 Drk11@1.0 Drk12@1.1 Drk13@1.2 Drk14@1.3 Drk15@1.4 Drk16@1.5

Drk17@1.6 Drk18@1.7;

i4 s4 q4 | OD2@.1 OD5@.4 OD6@.5 OD7@.6 OD8@.7 OD9@.8

OD10@.9 OD11@1.0 OD12@1.1 OD13@1.2 OD14@1.3 OD15@1.4 OD16@1.5

OD17@1.6;

!Note that for omitted indicators, the relative weights (corresponding to time) for the latent growth curves are adjusted accordingly (e.g. OD2@.1 OD5@.4).

[MJ1$1-MJ18$1@0];

[OD2$1@0 OD5$1-OD17$1@0];

[Drk1$1-Drk18$1@0];

[puff2$1-puff18$1@0];

%c#1%

[i1 s1 q1] (a1-a3);

[i2 s2 q2] (b1-b3);

[i3 s3 q3] (c1-c3);

[i4 s4 q4] (d1-d3);

!These commands (under %c#1%) instruct Mplus to independently estimate the means for these latent growth curve factors).

%c#2%

[i1 s1 q1] (a4-a6);

[i2 s2 q2] (b4-b6);

[i3 s3 q3] (c4-c6);

[i4 s4 q4] (d4-d6);

%c#3%

[i1 s1 q1] (a7-a9);

[i2 s2 q2] (b7-b9);

[i3 s3 q3] (c7-c9);

[i4 s4 q4] (d7-d9);

!Dean, Cole, & Bauer suggest to constrain the model to prevent parameters assuming extreme (less than -15) values. One disadvantage of this constraint is that it prevents formal testing of relative model fit relative to k-1 class models (TECH11, TECH14).

model constraint:

a1>-15;

a2>-15;

a3>-15;

a4>-15;

a5>-15;

!etc…

d6>-15;

d7>-15;

d8>-15;

d9>-15;

# OUTPUT EX 6a

!The output shows class distribution for the optimal (3-class) solution based on unshown class !enumeration considerations (Vuong-Lo-Mendel tests, LRTs, bootstrapped LRTs, number of cases in !each class). These are unconditional models, meaning that we do not include predictors of class !membership. For inclusion of predictors of class membership (or differences in distal outcomes !based on class membership in latent class analyses, refer to Asparouhov & Muthen, 2014.

SUMMARY OF ANALYSIS

Number of groups 1

Number of observations 1023

Number of dependent variables 67

Number of independent variables 0

Number of continuous latent variables 12

Number of categorical latent variables 1

MODEL FIT INFORMATION

Number of Free Parameters 38

Loglikelihood

H0 Value -5108.806

H0 Scaling Correction Factor 1.0818

for MLR

Information Criteria

Akaike (AIC) 10293.612

Bayesian (BIC) 10480.971

Sample-Size Adjusted BIC 10360.279

(n\* = (n + 2) / 24)

FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASSES

BASED ON THE ESTIMATED MODEL

Latent

Classes

1 348.99281 0.34115

2 147.39413 0.14408

3 526.61305 0.51477

FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASSES

BASED ON ESTIMATED POSTERIOR PROBABILITIES

Latent

Classes

1 348.99281 0.34115

2 147.39414 0.14408

3 526.61306 0.51477

FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASSES

BASED ON THEIR MOST LIKELY LATENT CLASS MEMBERSHIP

Class Counts and Proportions

Latent

Classes

1 295 0.28837

2 146 0.14272

3 582 0.56891

CLASSIFICATION QUALITY

Entropy 0.718

Average Latent Class Probabilities for Most Likely Latent Class Membership (Row)

by Latent Class (Column)

1 2 3

1 0.877 0.050 0.073

2 0.107 0.865 0.028

3 0.128 0.011 0.861

Classification Probabilities for the Most Likely Latent Class Membership (Column)

by Latent Class (Row)

1 2 3

1 0.741 0.045 0.214

2 0.100 0.857 0.043

3 0.041 0.008 0.951

Logits for the Classification Probabilities for the Most Likely Latent Class Membership (Column)

by Latent Class (Row)

1 2 3

1 1.241 -1.564 0.000

2 0.852 2.998 0.000

3 -3.143 -4.824 0.000

MODEL RESULTS

Two-Tailed

Estimate S.E. Est./S.E. P-Value

Latent Class 1

I1 |

MJ1 1.000 0.000 999.000 999.000

MJ2 1.000 0.000 999.000 999.000

MJ3 1.000 0.000 999.000 999.000

MJ4 1.000 0.000 999.000 999.000

MJ5 1.000 0.000 999.000 999.000

MJ6 1.000 0.000 999.000 999.000

MJ7 1.000 0.000 999.000 999.000

MJ8 1.000 0.000 999.000 999.000

MJ9 1.000 0.000 999.000 999.000

MJ10 1.000 0.000 999.000 999.000

MJ11 1.000 0.000 999.000 999.000

MJ12 1.000 0.000 999.000 999.000

MJ13 1.000 0.000 999.000 999.000

MJ14 1.000 0.000 999.000 999.000

MJ15 1.000 0.000 999.000 999.000

MJ16 1.000 0.000 999.000 999.000

MJ17 1.000 0.000 999.000 999.000

MJ18 1.000 0.000 999.000 999.000

S1 |

MJ1 0.000 0.000 999.000 999.000

MJ2 0.100 0.000 999.000 999.000

MJ3 0.200 0.000 999.000 999.000

MJ4 0.300 0.000 999.000 999.000

MJ5 0.400 0.000 999.000 999.000

MJ6 0.500 0.000 999.000 999.000

MJ7 0.600 0.000 999.000 999.000

MJ8 0.700 0.000 999.000 999.000

MJ9 0.800 0.000 999.000 999.000

MJ10 0.900 0.000 999.000 999.000

MJ11 1.000 0.000 999.000 999.000

MJ12 1.100 0.000 999.000 999.000

MJ13 1.200 0.000 999.000 999.000

MJ14 1.300 0.000 999.000 999.000

MJ15 1.400 0.000 999.000 999.000

MJ16 1.500 0.000 999.000 999.000

MJ17 1.600 0.000 999.000 999.000

MJ18 1.700 0.000 999.000 999.000

\*Note the weights of the quadratic component.

Q1 |

MJ1 0.000 0.000 999.000 999.000

MJ2 0.010 0.000 999.000 999.000

MJ3 0.040 0.000 999.000 999.000

MJ4 0.090 0.000 999.000 999.000

MJ5 0.160 0.000 999.000 999.000

MJ6 0.250 0.000 999.000 999.000

MJ7 0.360 0.000 999.000 999.000

MJ8 0.490 0.000 999.000 999.000

MJ9 0.640 0.000 999.000 999.000

MJ10 0.810 0.000 999.000 999.000

MJ11 1.000 0.000 999.000 999.000

MJ12 1.210 0.000 999.000 999.000

MJ13 1.440 0.000 999.000 999.000

MJ14 1.690 0.000 999.000 999.000

MJ15 1.960 0.000 999.000 999.000

MJ16 2.250 0.000 999.000 999.000

MJ17 2.560 0.000 999.000 999.000

MJ18 2.890 0.000 999.000 999.000

Means

!Substance-specific means on the logit scale are presented. The estimates can be transformed based on their relative contribution to the hazard function based on the time point (for instance, at time point 0, only the intercept is applied. At time point 1, the slope is applied at a factor of .1, and the quadratic effect is applied at a factor of .01 (equal to above), et cetera. Thus, patterns of initiation can be graphed per class and compared between classes. This yields something like the following:

*Figure note. The y-axis denotes the proportion of members (“one minus cumulative hazard”) of each latent class that remain substance-naïve for the specific substance: mj (Marijuana), puff (Cigarette puffing), and drk (Consuming a full drink of alcohol), at a given age.*

I1 -5.576 0.789 -7.063 0.000

S1 2.703 1.538 1.758 0.079

Q1 0.441 0.752 0.587 0.557

I2 -6.442 1.248 -5.160 0.000

S2 4.950 2.275 2.175 0.030

Q2 -1.660 1.044 -1.590 0.112

I3 -7.941 0.901 -8.812 0.000

S3 8.781 1.724 5.092 0.000

Q3 -2.597 0.814 -3.188 0.001

I4 -15.000 0.000 \*\*\*\*\*\*\*\*\* 0.000

S4 15.009 2.102 7.141 0.000

Q4 -5.263 1.429 -3.683 0.000

Thresholds

!All residual thresholds for indicators are constrained to be equal to zero. Omitted from below: Thresholds for other substances.

MJ1$1 0.000 0.000 999.000 999.000

MJ2$1 0.000 0.000 999.000 999.000

MJ3$1 0.000 0.000 999.000 999.000

MJ4$1 0.000 0.000 999.000 999.000

MJ5$1 0.000 0.000 999.000 999.000

MJ6$1 0.000 0.000 999.000 999.000

MJ7$1 0.000 0.000 999.000 999.000

MJ8$1 0.000 0.000 999.000 999.000

MJ9$1 0.000 0.000 999.000 999.000

MJ10$1 0.000 0.000 999.000 999.000

MJ11$1 0.000 0.000 999.000 999.000

MJ12$1 0.000 0.000 999.000 999.000

MJ13$1 0.000 0.000 999.000 999.000

MJ14$1 0.000 0.000 999.000 999.000

MJ15$1 0.000 0.000 999.000 999.000

MJ16$1 0.000 0.000 999.000 999.000

MJ17$1 0.000 0.000 999.000 999.000

MJ18$1 0.000 0.000 999.000 999.000

Latent Class 2

Means

I1 -7.630 1.473 -5.180 0.000

S1 14.986 4.504 3.327 0.001

Q1 -8.037 3.043 -2.641 0.008

I2 -5.456 0.782 -6.977 0.000

S2 11.511 2.811 4.095 0.000

Q2 -6.775 1.991 -3.402 0.001

I3 -6.502 0.880 -7.387 0.000

S3 13.070 3.206 4.077 0.000

Q3 -7.301 2.410 -3.030 0.002

I4 -6.581 1.420 -4.634 0.000

S4 7.442 3.013 2.470 0.014

Q4 -3.573 1.538 -2.322 0.020

Latent Class 3

Means

I1 -6.823 0.628 -10.869 0.000

S1 12.741 1.329 9.587 0.000

Q1 -15.000 0.000 \*\*\*\*\*\*\*\*\* 0.000

I2 -5.767 0.625 -9.233 0.000

S2 0.660 2.024 0.326 0.744

Q2 -0.268 1.358 -0.197 0.844

I3 -7.419 0.711 -10.431 0.000

S3 5.536 1.401 3.950 0.000

Q3 -1.493 0.653 -2.286 0.022

I4 -15.000 0.000 \*\*\*\*\*\*\*\*\* 0.000

S4 16.665 1.654 10.075 0.000

Q4 -8.847 0.864 -10.237 0.000

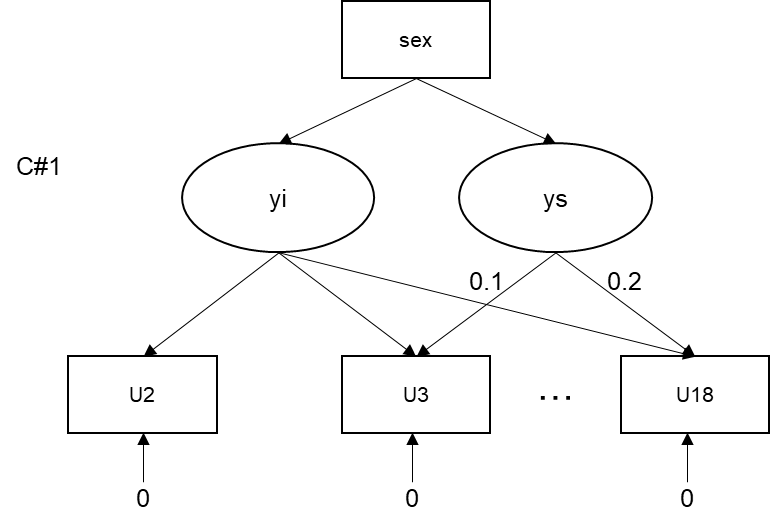
Categorical Latent Variables

Means

C#1 -0.411 0.125 -3.288 0.001

C#2 -1.273 0.209 -6.096 0.000

# INPUT EX 6b: Test of Proportional Hazards Assumption, Revisited



TITLE: EX 6b: MEPSUM single-indicator, single-class parallel to linear violation of PH

DATA:

FILE IS 20160416\_MEPSUM.dat;

VARIABLE:

NAMES ARE ID

MJ1-MJ26

DRK1-DRK26

PUFF1-PUFF26

OD1-OD26

sex hisp white cohort

;

USEVARIABLES =

MJ1 - MJ18 sex

;

IDvariable = ID;

Missing=all(999);

CLASSES = C(1);

CATEGORICAL =

MJ1 - MJ18

;

Analysis:

TYPE = Mixture;

process=4;

coverage=0;

model:

%OVERALL%

i1 s1 q1 | MJ1@.0 MJ2@.1 MJ3@.2 MJ4@.3 MJ5@.4 MJ6@.5 MJ7@.6 MJ8@.7 MJ9@.8

MJ10@.9 MJ11@1.0 MJ12@1.1 MJ13@1.2 MJ14@1.3 MJ15@1.4 MJ16@1.5

MJ17@1.6 MJ18@1.7;

!by regressing the slope representing linear change in hazard over time, and the quadratic slope representing quadratic change in hazard over time, on sex, we obtain estimates of the violation of proportional hazards in the effect of sex on initiation.

i1 s1 q1 on sex;

[MJ1$1-MJ18$1@0];

%c#1%

[i1 s1 q1] (a1-a3);

# OUTPUT EX 6b

SUMMARY OF ANALYSIS

Number of groups 1

Number of observations 1023

Number of dependent variables 18

Number of independent variables 1

Number of continuous latent variables 2

Number of categorical latent variables 1

MODEL FIT INFORMATION

Number of Free Parameters 4

Loglikelihood

H0 Value -1742.977

H0 Scaling Correction Factor 0.8951

for MLR

Information Criteria

Akaike (AIC) 3493.954

Bayesian (BIC) 3513.676

Sample-Size Adjusted BIC 3500.971

(n\* = (n + 2) / 24)

Latent Class 1

I1 |

MJ1 1.000 0.000 999.000 999.000

MJ2 1.000 0.000 999.000 999.000

MJ3 1.000 0.000 999.000 999.000

MJ4 1.000 0.000 999.000 999.000

MJ5 1.000 0.000 999.000 999.000

MJ6 1.000 0.000 999.000 999.000

MJ7 1.000 0.000 999.000 999.000

MJ8 1.000 0.000 999.000 999.000

MJ9 1.000 0.000 999.000 999.000

MJ10 1.000 0.000 999.000 999.000

MJ11 1.000 0.000 999.000 999.000

MJ12 1.000 0.000 999.000 999.000

MJ13 1.000 0.000 999.000 999.000

MJ14 1.000 0.000 999.000 999.000

MJ15 1.000 0.000 999.000 999.000

MJ16 1.000 0.000 999.000 999.000

MJ17 1.000 0.000 999.000 999.000

MJ18 1.000 0.000 999.000 999.000

S1 |

MJ1 0.000 0.000 999.000 999.000

MJ2 0.100 0.000 999.000 999.000

MJ3 0.200 0.000 999.000 999.000

MJ4 0.300 0.000 999.000 999.000

MJ5 0.400 0.000 999.000 999.000

MJ6 0.500 0.000 999.000 999.000

MJ7 0.600 0.000 999.000 999.000

MJ8 0.700 0.000 999.000 999.000

MJ9 0.800 0.000 999.000 999.000

MJ10 0.900 0.000 999.000 999.000

MJ11 1.000 0.000 999.000 999.000

MJ12 1.100 0.000 999.000 999.000

MJ13 1.200 0.000 999.000 999.000

MJ14 1.300 0.000 999.000 999.000

MJ15 1.400 0.000 999.000 999.000

MJ16 1.500 0.000 999.000 999.000

MJ17 1.600 0.000 999.000 999.000

MJ18 1.700 0.000 999.000 999.000

I1 ON

SEX -0.302 0.208 -1.450 0.147

S1 ON

SEX 0.083 0.197 0.420 0.674

Intercepts

I1 -4.412 0.124 -35.569 0.000

S1 1.331 0.119 11.184 0.000

Thresholds

MJ1$1 0.000 0.000 999.000 999.000

MJ2$1 0.000 0.000 999.000 999.000

MJ3$1 0.000 0.000 999.000 999.000

MJ4$1 0.000 0.000 999.000 999.000

MJ5$1 0.000 0.000 999.000 999.000

MJ6$1 0.000 0.000 999.000 999.000

MJ7$1 0.000 0.000 999.000 999.000

MJ8$1 0.000 0.000 999.000 999.000

MJ9$1 0.000 0.000 999.000 999.000

MJ10$1 0.000 0.000 999.000 999.000

MJ11$1 0.000 0.000 999.000 999.000

MJ12$1 0.000 0.000 999.000 999.000

MJ13$1 0.000 0.000 999.000 999.000

MJ14$1 0.000 0.000 999.000 999.000

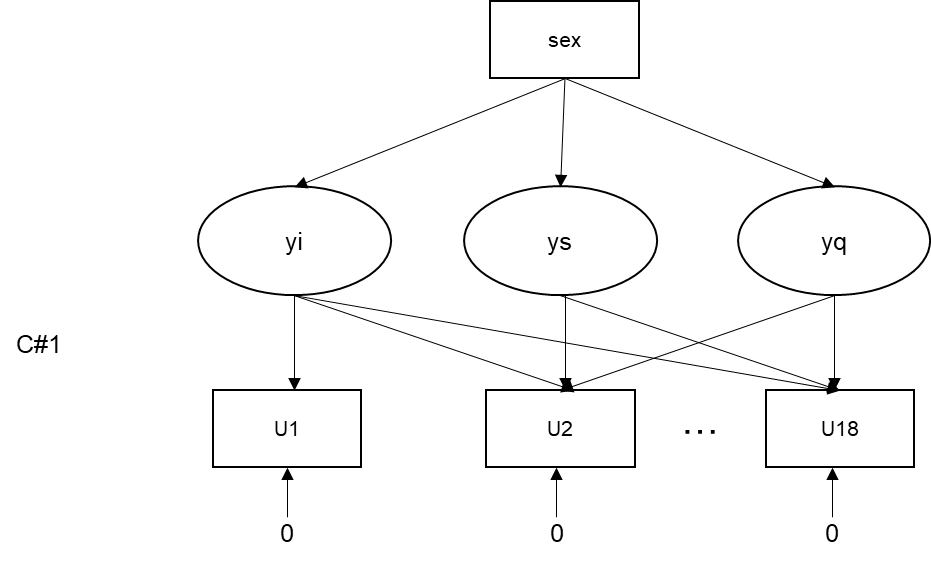
MJ15$1 0.000 0.000 999.000 999.000

MJ16$1 0.000 0.000 999.000 999.000

MJ17$1 0.000 0.000 999.000 999.000

MJ18$1 0.000 0.000 999.000 999.000

# INPUT EX 6c: Quadratic Proportional Hazards Violation



TITLE: EX 6c: a MEPSUM single-indicator, single-class parallel to quadratic violation of PH

DATA:

FILE IS 20160416\_MEPSUM.dat;

VARIABLE:

NAMES ARE ID

MJ1-MJ26

DRK1-DRK26

PUFF1-PUFF26

OD1-OD26

sex hisp white cohort

;

USEVARIABLES =

MJ1 - MJ18 sex

;

IDvariable = ID;

Missing=all(999);

CLASSES = C(1);

CATEGORICAL =

MJ1 - MJ18

;

Analysis:

TYPE = Mixture;

process=4;

coverage=0;

model:

%OVERALL%

i1 s1 | MJ1@.0 MJ2@.1 MJ3@.2 MJ4@.3 MJ5@.4 MJ6@.5 MJ7@.6 MJ8@.7 MJ9@.8

MJ10@.9 MJ11@1.0 MJ12@1.1 MJ13@1.2 MJ14@1.3 MJ15@1.4 MJ16@1.5

MJ17@1.6 MJ18@1.7;

i1 s1 on sex;

[MJ1$1-MJ18$1@0];

%c#1%

[i1 s1] (a1-a2);

# OUTPUT EX 6c

SUMMARY OF ANALYSIS

Number of groups 1

Number of observations 1023

Number of dependent variables 18

Number of independent variables 1

Number of continuous latent variables 3

Number of categorical latent variables 1

MODEL FIT INFORMATION

Number of Free Parameters 6

Loglikelihood

H0 Value -1724.471

H0 Scaling Correction Factor 0.9563

for MLR

Information Criteria

Akaike (AIC) 3460.942

Bayesian (BIC) 3490.525

Sample-Size Adjusted BIC 3471.468

(n\* = (n + 2) / 24)

Two-Tailed

Estimate S.E. Est./S.E. P-Value

Latent Class 1

I1 |

MJ1 1.000 0.000 999.000 999.000

MJ2 1.000 0.000 999.000 999.000

MJ3 1.000 0.000 999.000 999.000

MJ4 1.000 0.000 999.000 999.000

MJ5 1.000 0.000 999.000 999.000

MJ6 1.000 0.000 999.000 999.000

MJ7 1.000 0.000 999.000 999.000

MJ8 1.000 0.000 999.000 999.000

MJ9 1.000 0.000 999.000 999.000

MJ10 1.000 0.000 999.000 999.000

MJ11 1.000 0.000 999.000 999.000

MJ12 1.000 0.000 999.000 999.000

MJ13 1.000 0.000 999.000 999.000

MJ14 1.000 0.000 999.000 999.000

MJ15 1.000 0.000 999.000 999.000

MJ16 1.000 0.000 999.000 999.000

MJ17 1.000 0.000 999.000 999.000

MJ18 1.000 0.000 999.000 999.000

S1 |

MJ1 0.000 0.000 999.000 999.000

MJ2 0.100 0.000 999.000 999.000

MJ3 0.200 0.000 999.000 999.000

MJ4 0.300 0.000 999.000 999.000

MJ5 0.400 0.000 999.000 999.000

MJ6 0.500 0.000 999.000 999.000

MJ7 0.600 0.000 999.000 999.000

MJ8 0.700 0.000 999.000 999.000

MJ9 0.800 0.000 999.000 999.000

MJ10 0.900 0.000 999.000 999.000

MJ11 1.000 0.000 999.000 999.000

MJ12 1.100 0.000 999.000 999.000

MJ13 1.200 0.000 999.000 999.000

MJ14 1.300 0.000 999.000 999.000

MJ15 1.400 0.000 999.000 999.000

MJ16 1.500 0.000 999.000 999.000

MJ17 1.600 0.000 999.000 999.000

MJ18 1.700 0.000 999.000 999.000

Q1 |

MJ1 0.000 0.000 999.000 999.000

MJ2 0.010 0.000 999.000 999.000

MJ3 0.040 0.000 999.000 999.000

MJ4 0.090 0.000 999.000 999.000

MJ5 0.160 0.000 999.000 999.000

MJ6 0.250 0.000 999.000 999.000

MJ7 0.360 0.000 999.000 999.000

MJ8 0.490 0.000 999.000 999.000

MJ9 0.640 0.000 999.000 999.000

MJ10 0.810 0.000 999.000 999.000

MJ11 1.000 0.000 999.000 999.000

MJ12 1.210 0.000 999.000 999.000

MJ13 1.440 0.000 999.000 999.000

MJ14 1.690 0.000 999.000 999.000

MJ15 1.960 0.000 999.000 999.000

MJ16 2.250 0.000 999.000 999.000

MJ17 2.560 0.000 999.000 999.000

MJ18 2.890 0.000 999.000 999.000

I1 ON

SEX 0.507 0.397 1.275 0.202

!The significance of the slope and quadratic loading on sex indicates that there is an exponential curve to the estimated effect of sex on time-specific hazards.

S1 ON

SEX -2.209 0.957 -2.308 0.021

Q1 ON

SEX 1.282 0.539 2.377 0.017

Intercepts

I1 -5.691 0.289 -19.671 0.000

S1 5.062 0.682 7.424 0.000

Q1 -2.100 0.378 -5.554 0.000

Thresholds

MJ1$1 0.000 0.000 999.000 999.000

MJ2$1 0.000 0.000 999.000 999.000

MJ3$1 0.000 0.000 999.000 999.000

MJ4$1 0.000 0.000 999.000 999.000

MJ5$1 0.000 0.000 999.000 999.000

MJ6$1 0.000 0.000 999.000 999.000

MJ7$1 0.000 0.000 999.000 999.000

MJ8$1 0.000 0.000 999.000 999.000

MJ9$1 0.000 0.000 999.000 999.000

MJ10$1 0.000 0.000 999.000 999.000

MJ11$1 0.000 0.000 999.000 999.000

MJ12$1 0.000 0.000 999.000 999.000

MJ13$1 0.000 0.000 999.000 999.000

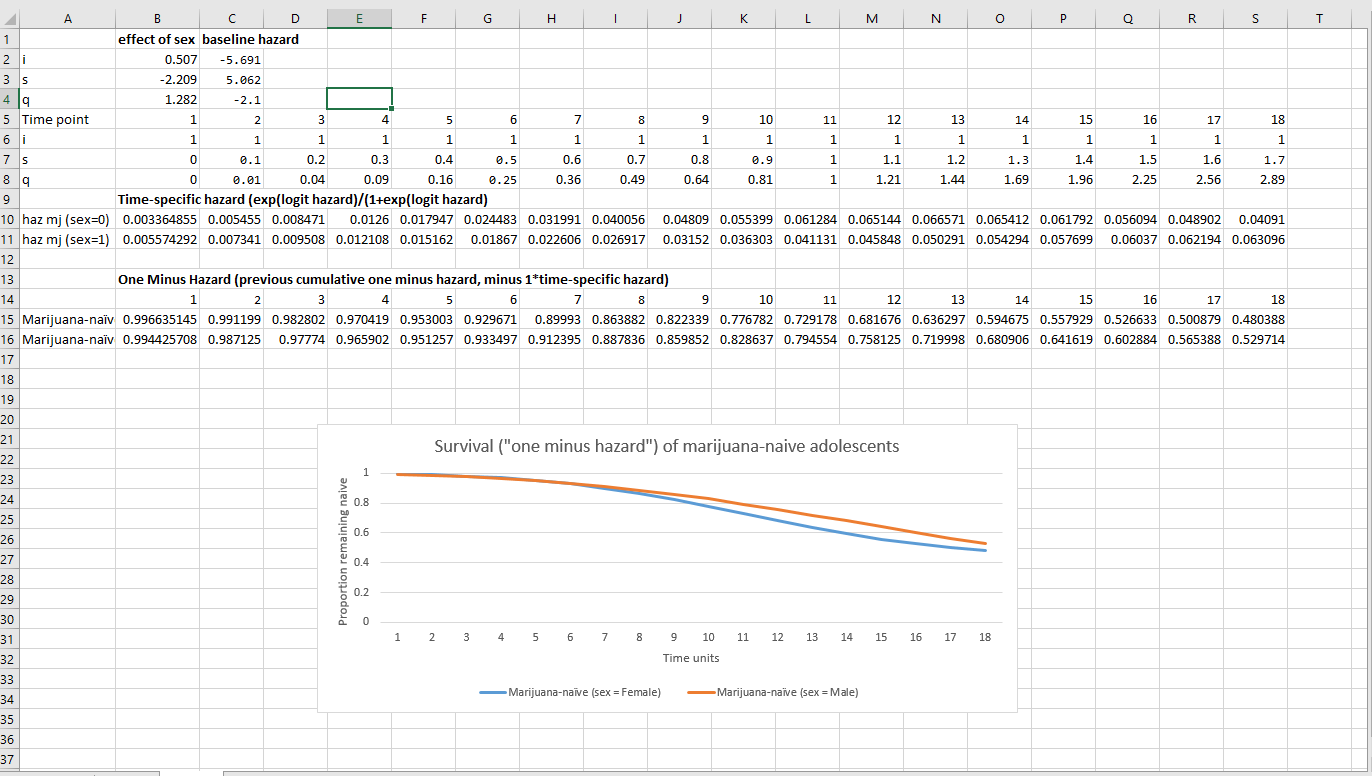
MJ14$1 0.000 0.000 999.000 999.000

MJ15$1 0.000 0.000 999.000 999.000

MJ16$1 0.000 0.000 999.000 999.000

MJ17$1 0.000 0.000 999.000 999.000

MJ18$1 0.000 0.000 999.000 999.000



*Excel allows us one way to graph the estimated survival over time for each value of the predictor sex. If the predictor were continuous, we could graph its effect at the mean, at -1 SD, and at +1 SD, for instance. See Singer & Willett, earlier cited page, for details on calculating the proportional hazard. Here, we need to take into account the relative contribution at each time point for the linear and quadratic effects on the cumulative hazard curve (see fixed loadings from model).*

# References:

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: Using the BCH method in Mplus to estimate a distal outcome model and an arbitrary secondary model. *Mplus Web Notes*, *21*(2), 1-22.

Dean, D. O., Cole, V., & Bauer, D. J. (2015). Delineating prototypical patterns of substance use initiations over time. *Addiction*, *110*(4), 585-594.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior research methods*, *40*(3), 879-891.

Singer, D., & Willett, J. B. (2003*). Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.

Yuan, K.H. and P.M. Bentler (1998). Robust mean and covariance structure analysis*. British journal of mathematical and statistical psychology, 51,* 63-88.