**Supplemental Appendix: SAS and R scripts for analyses**

**SAS Scripts**

/\*/ --------------------------------------------------------------------\\*\

| Terman Sample |

\\*\ --------------------------------------------------------------------/\*/

**data** Tvars;

filename Tvars 'c:\hawaii\T\_ph\_021011.dat';

infile tvars missover;

input id sex tener tprud tconf twill tcheer tmood tgrp tlead tpop texcel tsymp

tgen tconsc ttruth tknow torig tsense tintel ybirth pardiv educ tri50 smoker

alcnew midhlth illsumR bmi midadj age30 lastage status phys86 cvddth2 live70

live80 live85 include; **run**;

**proc** **means** data = tvars;

var tener tprud twill tcheer tmood tgrp tlead tpop texcel tsymp tgen tconsc

ttruth tknow torig tsense tintel;

**run**;

\* final FA with final vars;

**proc** **factor** data=tvars method=prinit nfact = **5** rotate=promax priors=smc

maxiter=**100** mineigen=**0.0** hey scree;

var Tener Tprud Twill Tcheer Tmood Tgrp Tlead Tpop Texcel Tsymp Tgen Tconsc

Ttruth Tknow Torig Tsense Tintel;

**proc** **corr** data = tvars alpha; var tprud tconsc ttruth texcel twill;

**proc** **corr** data = tvars alpha; var tsymp tgen;

**proc** **corr** data = tvars alpha; var tener tgrp tlead tpop tcheer;

**proc** **corr** data = tvars alpha; var tknow torig tintel tsense;

**proc** **corr** data = Tvars alpha; var midhlth illsumR midadj; **run**;

/\* Create a composite health variable \*/

**data** thlth1; set Tvars;

midhlthZ = midhlth;

illsumRz = illsumR;

midadjZ = midadj;

**proc** **standard** data = thlth1 out = thlth2 mean = **0** std = **1**;

var midhlthZ illsumRz midadjZ;

**data** Tpers; set Thlth2;

C = round(mean(tConsc,tPrud,tTruth,tWill,tExcel),**.1**);

E = round(mean(tEner,tGrp,tLead,tPop,tCheer),**.1**);

I = round(mean(tKnow,tOrig,tIntel,tSense),**.1**);

A = round(mean(tSymp,tGen),**.1**);

S = tmood;

hlth = (round(sum(midhlthZ,illsumRz,midadjZ),**.01**))+**9.2**;

**proc** **means** data = Tpers; var C E I A S hlth; **run**;

**proc** **corr** data = Tpers; var C E I A S; **run**;

**proc** **sort** data = Tpers; by sex;

**proc** **corr** data = Tpers; var C E I A S; by sex; **run**;

**proc** **corr** data = Tpers;

var C E I A S;

with sex ybirth pardiv educ tri50 smoker alcnew midhlth illsumR bmi midadj lastage status phys86 cvddth2 live70 live80 live85 hlth;

**run**;

**proc** **corr** data = Tpers; by sex;

var C E I A S;

with ybirth pardiv educ tri50 smoker alcnew midhlth illsumR bmi midadj lastage

status phys86 cvddth2 live70 live80 live85 hlth; **run**;

**proc** **reg** data = Tpers; model midhlth = C E I A S ybirth sex;

**proc** **reg** data = Tpers; model illsumR = C E I A S ybirth sex;

**proc** **reg** data = Tpers; model midadj = C E I A S ybirth sex;

**proc** **reg** data = Tpers; model hlth = C E I A S ybirth sex;

**proc** **reg** data = Tpers; model educ = C E I A S ybirth sex;

**proc** **reg** data = Tpers; model alcnew = C E I A S ybirth sex;

**run**; **quit**;

/\* test mediators \*/

**proc** **reg** data = Tpers; model midhlth = C E I A S age30 educ alcnew ybirth sex;

**proc** **reg** data = Tpers; model illsumR = C E I A S age30 educ alcnew ybirth sex;

**proc** **reg** data = Tpers; model midadj = C E I A S age30 educ alcnew ybirth sex;

**proc** **reg** data = Tpers; model hlth = C E I A S age30 educ alcnew ybirth sex;

**run**; **quit**;

**proc** **phreg** data = Tpers; model (age30,lastage)\*status(**0**) = C E I A S age30 ybirth sex/rl;

**proc** **phreg** data = Tpers; model (age30,lastage)\*status(**0**) = C E I A S age30 educ alcnew ybirth sex/rl;

**run**;

/\*/ --------------------------------------------------------------------\\*\

| Hawaii Sample |

\\*\ --------------------------------------------------------------------/\*/;

**data** Hbase;

filename hbase 'c:\hawaii\hi\_PH\_021011.dat';

infile hbase missover;

input id sex include consc fickleR impulseR irrespR pers planful energy lethargR

socconf assert greg submisR happy seclR curious orig imagine verbal consid

minimize spiteR rudeR acceptR touchyR jealousR fearR care carelssR ybirth divorce

educ triact smoker alcohol hlth illsumR bmi mental urine gluc cholHD tri sbp dbp

whr bmicl totchol bpmed cholmed dysctoff10 dysctoff25 dyscthigh10 dyscthigh25

Zdysreg;

**run**;

\* Final FA with included variables;

**proc** **factor** data=hbase method=prinit nfact=**5** rotate=promax priors=smc

maxiter=**100** mineigen=**0.0** hey scree;

var consc fickle impulse irresp pers planful consid minimize spite rude

energy letharg socconf assert greg submis happy secl curious orig

imagine verbal accept touchy jealous fear;

**run**;

**proc** **corr** data = hbase alpha; var consc fickleR irrespR pers planful care carelssR;

**proc** **corr** data = hbase alpha; var energy lethargR socconf assert greg submisR seclR;

**proc** **corr** data = hbase alpha; var curious orig imagine verbal;

**proc** **corr** data = hbase alpha; var consid minimize spiteR rudeR;

**proc** **corr** data = hbase alpha; var acceptR touchyR jealousR;

**proc** **corr** data = hbase alpha; var hlth illsumR mental;

**run**;

**data** hhlth1; set hbase;

hlthZ = hlth;

illsumRz = illsumR;

mentalZ = mental;

**proc** **standard** data = hhlth1 out = hhlth2 mean = **0** std = **1**;

var hlthZ illsumRz mentalZ;

**data** Hpers; set hhlth2;

C = round(mean(consc,fickleR,irrespR,pers,planful,care,carelssR),**.1**);

E = round(mean(energy,lethargR,socconf,assert,greg,submisR,seclR),**.1**);

I = round(mean(curious,orig,imagine,verbal),**.1**);

A = round(mean(consid,minimize,spiteR,rudeR),**.1**);

S = round(mean(acceptR,touchyR,jealousR),**.1**);

health = (round(sum(hlthZ,illsumRz,mentalZ),**.01**))+**9.66**;

**proc** **means** data = Hpers;

var C E I A S health; **run**;

**proc** **corr** data = Hpers; var C E I A S; **run**;

**proc** **sort** data = Hpers; by sex;

**proc** **corr** data = Hpers; var C E I A S; by sex; **run**;

**proc** **corr** data = Hpers;

var C E I A S;

with sex ybirth divorce educ triact smoker alcohol hlth illsumR bmi mental

urine gluc cholHD tri sbp dbp whr bmicl totchol bpmed cholmed dysctoff10

dysctoff25 dyscthigh10 dyscthigh25 Zdysreg health;

**run**;

**proc** **corr** data = Hpers; by sex;

var C E I A S;

with ybirth divorce educ triact smoker alcohol hlth illsumR bmi mental

urine gluc cholHD tri sbp dbp whr bmi totchol bpmed cholmed dysctoff10

dysctoff25 dyscthigh10 dyscthigh25 Zdysreg health; **run**;

**proc** **reg** data = Hvars2; model hlth = C E I A S ybirth sex;

**proc** **reg** data = Hvars2; model illsumR = C E I A S ybirth sex;

**proc** **reg** data = Hvars2; model mental = C E I A S ybirth sex;

**proc** **reg** data = Hvars2; model health = C E I A S ybirth sex;

**proc** **reg** data = Hvars2; model educ = C E I A S ybirth sex;

**proc** **reg** data = Hvars2; model alcohol = C E I A S ybirth sex;

**proc** **reg** data = Hvars2; model Zdysreg = C E I A S ybirth sex;

**run**; **quit**;

/\* include mediators \*/

**proc** **reg** data = Hvars2; model hlth = C E I A S educ alcohol ybirth sex;

**proc** **reg** data = Hvars2; model illsumR = C E I A S educ alcohol ybirth sex;

**proc** **reg** data = Hvars2; model mental = C E I A S educ alcohol ybirth sex;

**proc** **reg** data = Hvars2; model health = C E I A S educ alcohol ybirth sex;

**proc** **reg** data = Hvars2; model Zdysreg = C E I A S educ alcohol ybirth sex;

**run**; **quit**;

/\*/ --------------------------------------------------------------------\\*\

| Combining Samples |

\\*\ --------------------------------------------------------------------/\*/

\*\*\* standardize the personality traits;

**data** Tset; set Tpers;

id2 = id;

Cz = C;

Ez = E;

Iz = I;

Az = A;

Sz = S;

srh = midhlth;

mental = midadj;

alcohol = alcnew;

health = hlth;

group = **1**; \* grouping variable;

**proc** **standard** data = Tset out = Tz mean = **0** std = **1**;

var Cz Ez Iz Az Sz;

**proc** **means** data = Tz;

var srh mental alcohol illsumR health Cz Ez Iz Az Sz;

title 'combining datasets'; **run**;

**proc** **sort** data = Tz (keep = id2 sex srh mental alcohol educ illsumR health Cz Ez Iz Az Sz ybirth age30 group) out = Tz2; by id2; **run**;

**data** Hset; set Hpers;

id2 = id+**1500**;

Cz = C;

Ez = E;

Iz = I;

Az = A;

Sz = S;

srh = hlth;

group = **2**;

**proc** **standard** data = Hset mean = **0** std = **1** out = Hz;

var Cz Ez Iz Az Sz;

**proc** **means** data = Hz;

var srh mental alcohol illsumR health Cz Ez Iz Az Sz; **run**;

**proc** **sort** data = Hz (keep = id2 sex srh mental alcohol educ illsumR health Cz Ez Iz Az Sz ybirth group) out = Hz2; by id2; **run**;

**data** THcombo; merge Tz2 Hz2; by id2;

**proc** **means** data = THcombo;

var srh mental alcohol educ illsumR health Cz Ez Iz Az Sz; **run**;

**data** THcomb2; set THcombo;

Cz = round((Cz+**5.8478639**),**.001**);

Ez = round((Ez+**4.2303001**),**.001**);

Iz = round((Iz+**4.9435717**),**.001**);

Az = round((Az+**4.3138053**),**.001**);

Sz = round((Sz+**4.3905148**),**.001**);

illsum = **5** - illsumR; \* so 0 = none;

**proc** **means** data = THcomb2;

var Cz Ez Iz Az Sz; **run**;

/\* Create outfile for R

data Rset; set ages;

file 'c:\Hawaii\THcomb\_051311.dat';

put id2 Cz Ez Iz Az Sz sex srh mental illsumR health alcohol educ ybirth HIgrp;

run;

\*/

**R scripts**

**Terman Specific Models**

T1 <- read.table("Tpers062012.dat", sep='\t', header=TRUE, as.is = TRUE)

## check summary statistics and structure

summary(T1)

T1[1:9,]

str(T1)

library(lavaan)

Tpers <- '

# measuresment model

f1 =~ prud+cons+truth+excel+will

f2 =~ symp+gen

f3 =~ ener+grp+lead+pop+cheer

f4 =~ know+orig+intel+sense

f5 =~ mood

'

Tset1 <- cfa(Tpers, data = T1, missing = "ml")

summary(Tset1, fit.measure = TRUE)

## get standardized values

TEst <- parameterEstimates(Tset1, ci = TRUE, standardized = TRUE)

subset(TEst, op == "=~")

## Health model using full SEM model

Tph <- '

# measuresment model

f1 =~ prud+cons+truth+excel+will

f2 =~ symp+gen

f3 =~ ener+grp+lead+pop+cheer

f4 =~ know+orig+intel+sense

f5 =~ mood

f6 =~ midhlth+illsumR+midadj

# regression model

f6 ~ f1+f2+f3+f4+f5+ybirth+sex

'

Tpmh1 <- '

f1 =~ prud+cons+truth+excel+will

f2 =~ symp+gen

f3 =~ ener+grp+lead+pop+cheer

f4 =~ know+orig+intel+sense

f5 =~ mood

f6 =~ midhlth+illsumR+midadj

# regression model

f6 ~ f1+f2+f3+f4+f5+educ+alcnew

educ ~ f1+f2+f3+f4+f5+ybirth+sex

alcnew ~ f1+f2+f3+f4+f5+ybirth+sex

'

Tpmh2 <- '

f1 =~ prud+cons+truth+excel+will

f2 =~ symp+gen

f3 =~ ener+grp+lead+pop+cheer

f4 =~ know+orig+intel+sense

f5 =~ mood

f6 =~ midhlth+illsumR+midadj

# regression model

f6 ~ educ+alcnew

educ ~ f1+f2+f3+f4+f5+ybirth+sex

alcnew ~ f1+f2+f3+f4+f5+ybirth+sex

'

Tmodel1 <- sem(Tph, data = T1, missing = "ml")

Tmodel2 <- sem(Tpmh1, data = T1, missing = "ml")

Tmodel3 <- sem(Tpmh2, data = T1, missing = "ml")

summary(Tmodel1, fit.measure = TRUE)

summary(Tmodel2, fit.measure = TRUE)

summary(Tmodel3, fit.measure = TRUE)

## use model 3. Get paramaters

TmEst <- parameterEstimates(Tmodel3, ci = TRUE, standardized = TRUE)

subset(TmEst, op == "~")

anova(Tmodel2,Tmodel3)

# Self-rated health

Tsrh1 <- '

f1 =~ prud+cons+truth+excel+will

f2 =~ symp+gen

f3 =~ ener+grp+lead+pop+cheer

f4 =~ know+orig+intel+sense

f5 =~ mood

# regression model

midhlth ~ f1+f2+f3+f4+f5+educ+alcnew

educ ~ f1+f2+f3+f4+f5+ybirth+sex

alcnew ~ f1+f2+f3+f4+f5+ybirth+sex

'

Tsrh2 <- sem(Tsrh1, data = T1, missing = "ml")

summary(Tsrh2, fit.measure = TRUE)

Tsrh1b <- '

f1 =~ prud+cons+truth+excel+will

f2 =~ symp+gen

f3 =~ ener+grp+lead+pop+cheer

f4 =~ know+orig+intel+sense

f5 =~ mood

# regression model

midhlth ~ educ+alcnew

educ ~ f1+f2+f3+f4+f5+ybirth+sex

alcnew ~ f1+f2+f3+f4+f5+ybirth+sex

'

Tsrh2b <- sem(Tsrh1b, data = T1, missing = "ml")

summary(Tsrh2b, fit.measure = TRUE)

**Hawaii Specific Models**

H1 <- read.table("Hpers062012.dat", sep='\t', header=TRUE, as.is = TRUE)

## check summary statistics and structure

summary(H1)

H1[1:9,]

str(H1)

library(lavaan)

Hpers <- '

# measuresment model

f1 =~ cons+fickleR+irresR+persev+planful+care+carelsR

f2 =~ consid+mini+spiteR+rudeR

f3 =~ energy+lethargR+socconf+assert+greg+submisR+seclR

f4 =~ curious+orig+imagine+verbal

f5 =~ acceptR+touchyR+jealR

'

Hset1 <- cfa(Hpers, data = H1, missing = "ml")

summary(Hset1, fit.measure = TRUE)

# get parameter estimates

HEst <- parameterEstimates(Hset1, ci = TRUE, standardized = TRUE)

subset(HEst, op == "=~")

## Health model using full SEM model

Hph <- '

# measuresment model

f1 =~ cons+fickleR+irresR+persev+planful+care+carelsR

f2 =~ consid+mini+spiteR+rudeR

f3 =~ energy+lethargR+socconf+assert+greg+submisR+seclR

f4 =~ curious+orig+imagine+verbal

f5 =~ acceptR+touchyR+jealR

f6 =~ srh+illR+ment

# regression model

f6 ~ f1+f2+f3+f4+f5+ybirth+sex

'

Hpmh1 <- '

f1 =~ cons+fickleR+irresR+persev+planful+care+carelsR

f2 =~ consid+mini+spiteR+rudeR

f3 =~ energy+lethargR+socconf+assert+greg+submisR+seclR

f4 =~ curious+orig+imagine+verbal

f5 =~ acceptR+touchyR+jealR

f6 =~ srh+illR+ment

# regression model

f6 ~ f1+f2+f3+f4+f5+educ+alc

educ ~ f1+f2+f3+f4+f5+ybirth+sex

alc ~ f1+f2+f3+f4+f5+ybirth+sex

'

Hpmh2 <- '

f1 =~ cons+fickleR+irresR+persev+planful+care+carelsR

f2 =~ consid+mini+spiteR+rudeR

f3 =~ energy+lethargR+socconf+assert+greg+submisR+seclR

f4 =~ curious+orig+imagine+verbal

f5 =~ acceptR+touchyR+jealR

f6 =~ srh+illR+ment

# regression model

f6 ~ educ+alc

educ ~ f1+f2+f3+f4+f5+ybirth+sex

alc ~ f1+f2+f3+f4+f5+ybirth+sex

'

Hmodel1 <- sem(Hph, data = H1, missing = "ml")

Hmodel2 <- sem(Hpmh1, data = H1, missing = "ml")

Hmodel3 <- sem(Hpmh2, data = H1, missing = "ml")

summary(Hmodel1, fit.measure = TRUE)

summary(Hmodel2, fit.measure = TRUE)

summary(Hmodel3, fit.measure = TRUE)

## use model 3. Get paramaters

HmEst <- parameterEstimates(Hmodel3, ci = TRUE, standardized = TRUE)

subset(HmEst, op == "~")

anova(Hmodel2,Hmodel3)

# Just self-rated health

Hsrh1 <- '

f1 =~ cons+fickleR+irresR+persev+planful+care+carelsR

f2 =~ consid+mini+spiteR+rudeR

f3 =~ energy+lethargR+socconf+assert+greg+submisR+seclR

f4 =~ curious+orig+imagine+verbal

f5 =~ acceptR+touchyR+jealR

# regression model

srh ~ f1+f2+f3+f4+f5+educ+alc

educ ~ f1+f2+f3+f4+f5+ybirth+sex

alc ~ f1+f2+f3+f4+f5+ybirth+sex

'

Hsrh2 <- sem(Hsrh1, data = H1, missing = "ml")

summary(Hsrh2, fit.measure = TRUE)

**Combined Models**

set2 <- read.table("THset2.dat", sep='\t', header=TRUE, as.is = TRUE)

## check summary statistics and structure

summary(set2)

set2[1:9,]

str(set2)

cor(set2)

# attach the dataset

attach(set2)

### Now estimate SEM model using lavaan

library(lavaan)

## Multigroup model

grpModel <- '

# measuresment model

f1 =~ srh + illR + ment

# regressions

f1 ~ educ + alc

educ ~ con+ext+int+agr+neu+sex+ybirth

alc ~ con+ext+int+agr+neu+sex+ybirth

'

## include direct paths from personality to health

grp3Model <- '

# measuresment model

f1 =~ srh + illR + ment

# regressions

f1 ~ educ + alc +con+ext+int+agr+neu

educ ~ con+ext+int+agr+neu+sex+ybirth

alc ~ con+ext+int+agr+neu+sex+ybirth

'

## Multi-group model, with ml estimate for missing data

miss1 <- sem(grpModel, data = set2, group = "higrp", missing = "ml")

miss2 <- sem(grpModel, data = set2, group = "higrp", group.equal=c("loadings"), missing = "ml")

miss3 <- sem(grp3Model, data = set2, group = "higrp", missing = "ml")

miss4 <- sem(grp3Model, data = set2, group = "higrp", group.equal=c("loadings"), missing = "ml")

## to be complete, try setting everything to equal

miss5 <- sem(grpModel, data = set2, group = "higrp", group.equal=c("loadings", "intercepts", "means", "regressions", "residuals", "residual.covariances", "lv.variances", "lv.covariances"), missing = "ml")

miss6 <- sem(grp3Model, data = set2, group = "higrp", group.equal=c("loadings", "intercepts", "means", "regressions", "residuals", "residual.covariances", "lv.variances", "lv.covariances"), missing = "ml")

# get results and compare models

summary(miss1, fit.measure = TRUE)

summary(miss2, fit.measure = TRUE)

summary(miss3, fit.measure = TRUE)

summary(miss4, fit.measure = TRUE)

summary(miss5, fit.measure = TRUE)

summary(miss6, fit.measure = TRUE)

anova(miss1,miss2)

anova(miss3,miss4)

anova(miss2,miss4)

anova(miss5,miss6)

anova(miss2,miss5)

anova(miss3,miss6)

anova(miss1,miss5)

# the additional constraints are much worse

# for figure 4, use miss1. Get standardized values

MainEst <- parameterEstimates(miss1, ci = TRUE, standardized = TRUE)

subset(MainEst, op == c("=~","~"))

### Only self-rated health (so path model)

srhModel <- '

# regressions

srh ~ educ + alc +con+ext+int+agr+neu

educ ~ con+ext+int+agr+neu+sex+ybirth

alc ~ con+ext+int+agr+neu+sex+ybirth

'

srh1 <- sem(srhModel, data = set2, group = "higrp", missing = "ml")

summary(srh1, fit.measure = TRUE)

?sem()

srh2Model <- '

# regressions

srh ~ educ + alc

educ ~ con+ext+int+agr+neu+sex+ybirth

alc ~ con+ext+int+agr+neu+sex+ybirth

'

srh2 <- sem(srh2Model, data = set2, group = "higrp", missing = "ml")

summary(srh2, fit.measure = TRUE)

anova(srh1, srh2)

detach(set2)