Multivariate Statistics

Assignment 1

Wenhan Cu r0973145

Tom Leppens r0716859

# Question A:

To start this question, we first loaded the ESS data in R. Then we centered the data and fitted the model.

|  |
| --- |
| #compute centered data  cess<-ess  cess[,2:14]<-scale(ess[,2:14],center=TRUE,scale=FALSE)    #we impose constraints on factors and accept correlation between factors  cfa1<- '  social\_trust =~NA\*ppltrst+pplfair+pplhlp  trust\_institutions =~NA\*trstprl+trstlgl+trstplc+trstplt  well\_being =~NA\*fltdpr+fltsd+fltanx+wrhpp+enjlf+fltpcfl  social\_trust~~1\*social\_trust  trust\_institutions~~1\*trust\_institutions  well\_being~~1\*well\_being  social\_trust~~trust\_institutions  social\_trust~~well\_being  trust\_institutions~~well\_being  '  covmat <- cov(cess[,2:14])  fitcfa1 <- cfa(cfa1, sample.cov = covmat, sample.nobs=4046)  summary(fitcfa1, fit.measures=TRUE)  standardizedSolution(fitcfa1) |

|  |
| --- |
| > summary(fitcfa1, fit.measures=TRUE)  Latent Variables:  Estimate Std.Err z-value P(>|z|)  social\_trust =~  ppltrst 1.475 0.037 40.013 0.000  pplfair 1.285 0.034 37.927 0.000  pplhlp 1.313 0.036 36.620 0.000  trust\_institutions =~  trstprl 1.846 0.034 55.035 0.000  trstlgl 1.677 0.034 48.664 0.000  trstplc 1.301 0.035 37.324 0.000  trstplt 1.742 0.031 56.210 0.000  well\_being =~  fltdpr 0.466 0.011 43.278 0.000  fltsd 0.454 0.010 43.975 0.000  fltanx 0.450 0.012 37.467 0.000  wrhpp 0.570 0.012 48.097 0.000  enjlf 0.549 0.012 44.555 0.000  fltpcfl 0.514 0.014 37.934 0.000 |

Our interpretation: Loadings are all positive and significant. But factors social\_trust and trust\_instttns have positive moderate significant correlations.

|  |
| --- |
| > standardizedSolution(fitcfa1)  lhs op rhs est.std se z pvalue ci.lower ci.upper  1 social\_trust =~ ppltrst 0.684 0.013 52.036 0 0.658 0.709  2 social\_trust =~ pplfair 0.648 0.013 48.322 0 0.622 0.674  3 social\_trust =~ pplhlp 0.626 0.014 46.031 0 0.600 0.653  4 trust\_institutions =~ trstprl 0.789 0.008 93.956 0 0.773 0.805  5 trust\_institutions =~ trstlgl 0.718 0.010 74.774 0 0.699 0.737  6 trust\_institutions =~ trstplc 0.581 0.012 48.194 0 0.557 0.605  7 trust\_institutions =~ trstplt 0.802 0.008 97.758 0 0.786 0.818  8 well\_being =~ fltdpr 0.661 0.011 60.710 0 0.640 0.683  9 well\_being =~ fltsd 0.670 0.011 62.343 0 0.649 0.691  10 well\_being =~ fltanx 0.589 0.012 48.379 0 0.565 0.612  11 well\_being =~ wrhpp 0.718 0.010 72.725 0 0.699 0.738  12 well\_being =~ enjlf 0.677 0.011 63.729 0 0.656 0.698  13 well\_being =~ fltpcfl 0.595 0.012 49.291 0 0.571 0.618  14 social\_trust ~~ social\_trust 1.000 0.000 NA NA 1.000 1.000  15 trust\_institutions ~~ trust\_institutions 1.000 0.000 NA NA 1.000 1.000  16 well\_being ~~ well\_being 1.000 0.000 NA NA 1.000 1.000  17 social\_trust ~~ trust\_institutions 0.555 0.016 34.183 0 0.524 0.587  18 social\_trust ~~ well\_being 0.287 0.020 14.604 0 0.248 0.326  19 trust\_institutions ~~ well\_being 0.185 0.018 10.022 0 0.149 0.221  ... |

* convergent validity, divergent validity

Factor loadings highlighted (e.g. ppltrst, trstprl, trstlgl, trstplt, wrhpp has strong convergent validity) indicates a strong relationship between factor and variables and are statistically significant if 0.68 is our benchmark.

We can say a lack of divergent validity between social trust and trust\_institutions due to the high covariance 0.555.

* composite reliability of the factor scores

|  |
| --- |
| > reliability\_social\_trust  [1] 0.467856 0.419904 0.391876  Ppltrst Pplfair pplhlp |

Only a small part of variance is due to laten t variable social\_trust. Not a good measure.

|  |
| --- |
| > reliability\_trust\_institutions  [1] 0.622521 0.515524 0.337561 0.643204  Trstprl Trstlgl Trstplc trstplt |

Latent factor trust\_institutions is a good masure of Trstprl and trstplt.

|  |
| --- |
| > reliability\_well\_being  [1] 0.436921 0.448900 0.346921 0.515524 0.458329 0.354025  Fltdpr Fltsd Fltanx Wrhpp Enjlf Fltpcfl |

Only a small part of variance is due to laten t variable well\_being. Not a good measure.

|  |
| --- |
| > composite\_reliability\_social\_trust [1] 0.4264549  > composite\_reliability\_trust\_institutions [1] 0.5297277  > composite\_reliability\_well\_being [1] 0.426797 |

Reliability of factor score social\_trust, trust\_institution and well\_being are questionable.

* fit measures of the fitted CFA model

|  |
| --- |
| > fitmeasures(fitcfa1,c("cfi","tli","rmsea","srmr"))  cfi tli rmsea srmr  0.912 0.889 0.076 0.040 |

TLI is smaller than 0.95, CFI is close to 0.95 but still small, which mean there is room to improve the model. rmsea and srmr are smaller then 0.08 (cutoff<0.08) is satisfied.

## Question B (to be continued)

|  |
| --- |
| > modificationindices(fitcfa1)  lhs op rhs mi epc sepc.lv sepc.all sepc.nox  33 social\_trust =~ trstprl 15.852 -0.181 -0.181 -0.077 -0.077  35 social\_trust =~ trstplc 35.878 0.288 0.288 0.129 0.129  43 trust\_institutions =~ ppltrst 20.844 0.242 0.242 0.112 0.112  44 trust\_institutions =~ pplfair 19.103 -0.206 -0.206 -0.104 -0.104  58 well\_being =~ trstplt 16.240 -0.117 -0.117 -0.054 -0.054  ...  80 pplfair ~~ enjlf 29.089 0.093 0.093 0.103 0.103  82 pplhlp ~~ trstprl 15.508 -0.190 -0.190 -0.081 -0.081  84 pplhlp ~~ trstplc 30.187 0.297 0.297 0.100 0.100  92 trstprl ~~ trstlgl 52.242 -0.492 -0.492 -0.210 -0.210  93 trstprl ~~ trstplc 211.717 -0.854 -0.854 -0.326 -0.326  94 trstprl ~~ trstplt 559.300 1.707 1.707 0.914 0.914  101 trstlgl ~~ trstplc 478.146 1.275 1.275 0.430 0.430  102 trstlgl ~~ trstplt 168.787 -0.834 -0.834 -0.395 -0.395  109 trstplc ~~ trstplt 73.909 -0.471 -0.471 -0.199 -0.199  122 fltdpr ~~ fltsd 153.517 0.068 0.068 0.255 0.255  123 fltdpr ~~ fltanx 50.754 0.045 0.045 0.137 0.137  124 fltdpr ~~ wrhpp 62.404 -0.051 -0.051 -0.173 -0.173  125 fltdpr ~~ enjlf 68.018 -0.054 -0.054 -0.171 -0.171  127 fltsd ~~ fltanx 128.133 0.068 0.068 0.219 0.219  128 fltsd ~~ wrhpp 57.439 -0.047 -0.047 -0.168 -0.168  129 fltsd ~~ enjlf 86.209 -0.058 -0.058 -0.194 -0.194  130 fltsd ~~ fltpcfl 23.523 -0.033 -0.033 -0.094 -0.094  131 fltanx ~~ wrhpp 114.875 -0.075 -0.075 -0.219 -0.219  132 fltanx ~~ enjlf 150.356 -0.088 -0.088 -0.239 -0.239  133 fltanx ~~ fltpcfl 33.002 0.045 0.045 0.105 0.105  134 wrhpp ~~ enjlf 612.318 0.182 0.182 0.553 0.553 |

We can see that trstplc also have component in social\_trust, not just trust\_institution; Ppltrst and pplfair also have component in trust\_institution; Trstplt also have component in well\_being.

Adding those relationships will improve chi square goodness of fit.

# Question C

# Question D

|  |
| --- |
| #standardize variables  sess<-ess  sess[,2:14]<-scale(ess[,2:14],center=TRUE,scale=TRUE)  head(sess)    #conduct canonical correlation analysis  cancor.sess <- cancor(cbind(fltdpr, fltsd, fltanx, wrhpp, enjlf, fltpcfl)  ~ppltrst+pplfair+pplhlp+trstprl+trstlgl+trstplc+trstplt,  data=sess)    #print summary result  summary(cancor.sess)    #print canonical loadings  cancor.sess$structure$X.xscores  cancor.sess$structure$Y.yscores    #print redundancies (this tells us how much variance in Y is explained by X)  redu <- redundancy(cancor.sess)  round(redu$Xcan.redun, 3)  round(redu$Ycan.redun, 3) |

|  |
| --- |
| > summary(cancor.sess)    Canonical correlation analysis of:  7 X variables: ppltrst, pplfair, pplhlp, trstprl, trstlgl, trstplc, trstplt  with 6 Y variables: fltdpr, fltsd, fltanx, wrhpp, enjlf, fltpcfl    CanR CanRSQ Eigen percent cum scree  1 0.242629 5.887e-02 6.255e-02 77.37503 77.38 \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  2 0.110279 1.216e-02 1.231e-02 15.22875 92.60 \*\*\*\*\*\*  3 0.063206 3.995e-03 4.011e-03 4.96159 97.57 \*\*  4 0.041142 1.693e-03 1.696e-03 2.09741 99.66 \*  5 0.016167 2.614e-04 2.614e-04 0.32339 99.99  6 0.003343 1.118e-05 1.118e-05 0.01383 100.00 |

It seems that u only accounts for only a small portion of variance in t.

Discuss the results of the canonical correlation analysis. How many canonical correlations are significant?

|  |
| --- |
| Test of H0: The canonical correlations in the  current row and all that follow are zero    CanR LR test stat approx F numDF denDF Pr(> F)  1 0.242629 0.92415 7.6396 42 18920 < 2.2e-16 \*\*\*  2 0.110279 0.98196 2.4539 30 16138 1.618e-05 \*\*\*  3 0.063206 0.99405 1.2056 20 13384 0.2378  4 0.041142 0.99804 0.6617 12 10678 0.7897  5 0.016167 0.99973 0.1834 6 8074 0.9815  6 0.003343 0.99999 NaN 2 NaN NaN  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1    Raw canonical coefficients    X variables:  Xcan1 Xcan2 Xcan3 Xcan4 Xcan5 Xcan6  ppltrst -0.23396 0.615795 0.23325 -0.106850 0.028413 -0.935860  pplfair -0.55952 -0.911243 -0.31731 0.279475 -0.149407 -0.123845  pplhlp -0.15850 0.587203 -0.54639 0.295785 0.431509 0.529875  trstprl -0.22870 -0.359367 0.75026 -0.208300 0.859583 -0.011333  trstlgl -0.18737 0.134048 0.23261 -0.021502 0.259886 0.612462  trstplc -0.18360 -0.013577 -0.28297 -0.994752 -0.468963 -0.027708  trstplt 0.14576 0.262135 0.11677 0.758998 -1.179758 0.156499    Y variables:  Ycan1 Ycan2 Ycan3 Ycan4 Ycan5 Ycan6  fltdpr -0.27044 0.050904 0.18846 1.16023 -0.173882 0.326530  fltsd -0.14178 -0.011388 0.77970 -0.65374 0.551761 0.538595  fltanx -0.24734 0.908601 -0.35278 -0.07931 0.232582 -0.636455  wrhpp -0.18054 -0.028104 -1.07711 -0.26704 -0.036627 0.778898  enjlf -0.37609 -0.762763 0.18565 0.10128 0.384816 -0.921927  fltpcfl -0.15805 -0.030212 0.34612 -0.36593 -1.064937 -0.050794 |

|  |
| --- |
| > #print canonical loadings  > cancor.ess$structure$X.xscores  Xcan1 Xcan2 Xcan3 Xcan4 Xcan5 Xcan6  ppltrst -0.6763363 4.724725e-01 0.13525466 0.05282290 -0.01131331 -0.5374706  pplfair -0.8416969 -3.845485e-01 -0.24453178 0.24728799 -0.07490891 -0.1232267  pplhlp -0.6092904 4.750760e-01 -0.38969676 0.24141612 0.20860056 0.2809934  trstprl -0.5441966 -8.041144e-05 0.71491848 0.03091828 0.08653009 0.2225929  trstlgl -0.5990731 1.842299e-01 0.38310315 -0.18269912 -0.08660209 0.4894876  trstplc -0.5543285 1.441351e-01 0.01696339 -0.64550485 -0.40052327 0.2577618  trstplt -0.4406747 2.170732e-01 0.49023591 0.29713029 -0.55893911 0.2729865  > cancor.ess$structure$Y.yscores  Ycan1 Ycan2 Ycan3 Ycan4 Ycan5 Ycan6  fltdpr -0.7396063 0.12576041 0.18270422 0.56931392 -0.043348111 0.27890610  fltsd -0.6871178 0.12922116 0.43606685 -0.29073751 0.322580432 0.36390976  fltanx -0.6772528 0.67065702 -0.07524753 -0.08650432 0.099240187 -0.26182813  wrhpp -0.7485314 -0.18422145 -0.51344794 -0.17741963 -0.005463235 0.33260617  enjlf -0.7927583 -0.49962698 -0.07182186 -0.04339612 0.138156486 -0.30948079  fltpcfl -0.6438757 0.01868394 0.17137384 -0.27206474 -0.693025695 -0.03744553 |

How much of the variance in the Y variables can be explained by the X variables?

|  |
| --- |
| > #print redundancies (this tells us how much variance in Y is explained by X)  > redu <- redundancy(cancor.ess)  > round(redu$Xcan.redun, 3)  Xcan1 Xcan2 Xcan3 Xcan4 Xcan5 Xcan6  0.023 0.001 0.001 0.000 0.000 0.000  > round(redu$Ycan.redun, 3)  Ycan1 Ycan2 Ycan3 Ycan4 Ycan5 Ycan6  0.030 0.002 0.000 0.000 0.000 0.000 |

This confirmatory factor analysis model has three correlated factors and assumes each item only has a loading on the concept it aims to measure. To get an idea of the fit of the model we used fit measures, the standardized solution, and computed, for each latent variable, the composite reliability (CR), the average variance extracted (AVE), and the maximum shared variance (MSV) with other latent variables.

To test the validity and reliability we used the following cutoff values:

* CFI > 0.95
* TLI > 0.95
* RMSEA < 0.08
* SRMR < 0.08
* Convergent Validity: AVE > 0.5
* Discriminant Validity: MSV < AVE
* Reliability: CR > 0.7

Fitting the model yields the following results:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Chisq | df | p-value | CFI | TLI | RMSEA | SRMR |
| 689.143 | 32 | 0.000 | 0.944 | 0.921 | 0.071 | 0.034 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Factor score** | **AVE** | **MSV** | **CR** |
| **social trust** | 0.4264263 | 0.4732259 | 0.6901061 |
| **trust institutions** | 0.5294827 | 0.6433114 | 0.8160278 |
| **well being** | 0.4873276 | 0.5577240 | 0.7396447 |

Based on the fit measures we can state the following: The chi-square test tells us that the model is significantly different from a perfectly fitted model (chi-square=689.143, df=32, p-value<0.000). However, as this dataset has a large sample size (n=4046) the goodness of fit test can be significantly different even if the difference between the two models is very small. Nonetheless, the other measures indicate that the model cfa1 can probably be improved. Both CFI (0.944) and TLI (0.921) are lower than 0.95 and the RMSEA (0.071) and SRMR (0.034) are higher than 0.08. Furthermore, the validity measurements indicate that there could be improvements: the MSV is not always lower than the AVE, the CR for social trust is not bigger than 0.7, and the AVE is not always above 0.5.

**Question B**: Use modification indices to see how you can obtain a model that meets the criteria of good fit in (in terms of TLI, CFI, RMSEA, SRMR) by including a few well-chosen correlated error terms for pairs of items. Try to justify the correlated error terms from a substantive point of view

To improve our model, we can use the ‘modificationIndices()’ function to get an idea of which error terms correlation we can add to improve our model.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **lhs** | **op** | **rhs** | **mi** | **epc** | **sepc.lv** | **sepc.all** | **sepc.nox** |
| social\_trust | =~ | trstprl | 16.004 | -0.182 | -0.182 | -0.078 | -0.078 |
| social\_trust | =~ | trstlgl | 11.091 | 0.155 | 0.155 | 0.066 | 0.066 |
| social\_trust | =~ | trstplc | 34.495 | 0.283 | 0.283 | 0.126 | 0.126 |
| social\_trust | =~ | trstplt | 7.946 | -0.119 | -0.119 | -0.055 | -0.055 |
| social\_trust | =~ | fltdpr | 1.150 | 0.013 | 0.013 | 0.019 | 0.019 |
| social\_trust | =~ | fltsd | 7.992 | -0.034 | -0.034 | -0.050 | -0.050 |
| social\_trust | =~ | fltanx | 4.218 | 0.027 | 0.027 | 0.036 | 0.036 |
| trust\_institutions | =~ | ppltrst | 18.110 | 0.228 | 0.228 | 0.106 | 0.106 |
| trust\_institutions | =~ | pplfair | 15.552 | -0.185 | -0.185 | -0.094 | -0.094 |
| trust\_institutions | =~ | pplhlp | 0.168 | -0.020 | -0.020 | -0.010 | -0.010 |
| trust\_institutions | =~ | fltdpr | 0.188 | 0.005 | 0.005 | 0.007 | 0.007 |
| trust\_institutions | =~ | fltsd | 0.723 | -0.009 | -0.009 | -0.013 | -0.013 |
| trust\_institutions | =~ | fltanx | 0.251 | 0.006 | 0.006 | 0.008 | 0.008 |
| well\_being | =~ | ppltrst | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| well\_being | =~ | pplfair | 0.590 | 0.028 | 0.028 | 0.014 | 0.014 |
| well\_being | =~ | pplhlp | 0.594 | -0.029 | -0.029 | -0.014 | -0.014 |
| well\_being | =~ | trstprl | 0.159 | -0.013 | -0.013 | -0.006 | -0.006 |
| well\_being | =~ | trstlgl | 5.796 | 0.083 | 0.083 | 0.035 | 0.035 |
| well\_being | =~ | trstplc | 5.440 | 0.084 | 0.084 | 0.037 | 0.037 |
| well\_being | =~ | trstplt | 9.847 | -0.095 | -0.095 | -0.044 | -0.044 |
| ppltrst | ~~ | pplfair | 0.488 | -0.059 | -0.059 | -0.025 | -0.025 |
| ppltrst | ~~ | pplhlp | 11.715 | -0.296 | -0.296 | -0.116 | -0.116 |
| ppltrst | ~~ | trstprl | 0.702 | 0.041 | 0.041 | 0.018 | 0.018 |
| ppltrst | ~~ | trstlgl | 2.060 | 0.074 | 0.074 | 0.029 | 0.029 |
| ppltrst | ~~ | trstplc | 0.378 | -0.033 | -0.033 | -0.012 | -0.012 |
| ppltrst | ~~ | trstplt | 1.744 | 0.059 | 0.059 | 0.029 | 0.029 |
| ppltrst | ~~ | fltdpr | 0.912 | -0.016 | -0.016 | -0.020 | -0.020 |
| ppltrst | ~~ | fltsd | 1.018 | -0.016 | -0.016 | -0.023 | -0.023 |
| ppltrst | ~~ | fltanx | 4.422 | 0.039 | 0.039 | 0.042 | 0.042 |
| pplfair | ~~ | pplhlp | 16.453 | 0.297 | 0.297 | 0.119 | 0.119 |
| pplfair | ~~ | trstprl | 2.782 | -0.076 | -0.076 | -0.035 | -0.035 |
| pplfair | ~~ | trstlgl | 1.289 | 0.055 | 0.055 | 0.022 | 0.022 |
| pplfair | ~~ | trstplc | 3.774 | 0.099 | 0.099 | 0.036 | 0.036 |
| pplfair | ~~ | trstplt | 12.011 | -0.145 | -0.145 | -0.073 | -0.073 |
| pplfair | ~~ | fltdpr | 6.819 | 0.041 | 0.041 | 0.053 | 0.053 |
| pplfair | ~~ | fltsd | 0.043 | 0.003 | 0.003 | 0.004 | 0.004 |
| pplfair | ~~ | fltanx | 4.821 | -0.038 | -0.038 | -0.043 | -0.043 |
| pplhlp | ~~ | trstprl | 15.802 | -0.192 | -0.192 | -0.082 | -0.082 |
| pplhlp | ~~ | trstlgl | 0.500 | 0.036 | 0.036 | 0.014 | 0.014 |
| pplhlp | ~~ | trstplc | 30.995 | 0.301 | 0.301 | 0.101 | 0.101 |
| pplhlp | ~~ | trstplt | 0.165 | -0.018 | -0.018 | -0.009 | -0.009 |
| pplhlp | ~~ | fltdpr | 0.132 | -0.006 | -0.006 | -0.007 | -0.007 |
| pplhlp | ~~ | fltsd | 6.440 | -0.040 | -0.040 | -0.054 | -0.054 |
| pplhlp | ~~ | fltanx | 7.170 | 0.049 | 0.049 | 0.052 | 0.052 |
| trstprl | ~~ | trstlgl | 51.686 | -0.489 | -0.489 | -0.209 | -0.209 |
| trstprl | ~~ | trstplc | 210.055 | -0.850 | -0.850 | -0.324 | -0.324 |
| Trstprl | ~~ | trstplt | 554.999 | 1.700 | 1.700 | 0.911 | 0.911 |
| Trstprl | ~~ | fltdpr | 0.156 | 0.006 | 0.006 | 0.008 | 0.008 |
| Trstprl | ~~ | fltsd | 2.284 | 0.022 | 0.022 | 0.034 | 0.034 |
| Trstprl | ~~ | fltanx | 3.768 | -0.033 | -0.033 | -0.039 | -0.039 |
| trstlgl | ~~ | trstplc | 479.201 | 1.277 | 1.277 | 0.430 | 0.430 |
| trstlgl | ~~ | trstplt | 169.431 | -0.835 | -0.835 | -0.395 | -0.395 |
| trstlgl | ~~ | fltdpr | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| trstlgl | ~~ | fltsd | 0.645 | 0.012 | 0.012 | 0.017 | 0.017 |
| trstlgl | ~~ | fltanx | 1.307 | 0.021 | 0.021 | 0.022 | 0.022 |
| Trstplc | ~~ | trstplt | 73.607 | -0.470 | -0.470 | -0.198 | -0.198 |
| Trstplc | ~~ | fltdpr | 1.134 | -0.018 | -0.018 | -0.020 | -0.020 |
| Trstplc | ~~ | fltsd | 0.586 | 0.013 | 0.013 | 0.015 | 0.015 |
| Trstplc | ~~ | fltanx | 3.155 | 0.034 | 0.034 | 0.032 | 0.032 |
| Trstplt | ~~ | fltdpr | 0.030 | 0.002 | 0.002 | 0.004 | 0.004 |
| Trstplt | ~~ | fltsd | 4.301 | -0.028 | -0.028 | -0.047 | -0.047 |
| Trstplt | ~~ | fltanx | 0.458 | -0.011 | -0.011 | -0.014 | -0.014 |
| fltdpr | ~~ | fltsd | 3.441 | 0.041 | 0.041 | 0.182 | 0.182 |
| Fltdpr | ~~ | fltanx | 6.709 | -0.049 | -0.049 | -0.167 | -0.167 |
| fltsd | ~~ | fltanx | 1.012 | 0.020 | 0.020 | 0.077 | 0.077 |

Based on the output we included the following every proposal that lowers the chi square by at least 50 which yields the following model:

cfa2 <-' social\_trust =~ NA\*ppltrst + pplfair + pplhlp

trust\_institutions =~ NA\*trstprl + trstlgl + trstplc + trstplt

well\_being =~ NA\*fltdpr + fltsd + fltanx

social\_trust ~~ 1\*social\_trust

trust\_institutions ~~ 1\*trust\_institutions

well\_being ~~ 1\*well\_being

trstprl ~~ trstplc

trstlgl ~~ trstplc

trstlgl ~~ trstplt

trstplc ~~ trstplt

trstprl ~~ trstplt

trstprl ~~ trstlgl

'

Looking at the fit measures we see that the goodness of fit test is still significantly different (chi-square=85.457, df=27, p-value<0.000) but CFI and TLI are both above 0.95 and RMSEA and SRMR have decreased. However, they are still not below the cutoff of 0.08. Overall we can conclude that the model is improved and fits the data better.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Chisq | df | p-value | CFI | TLI | RMSEA | SRMR |
| 85.457 | 26 | 0.000 | 0.995 | 0.991 | 0.024 | 0.014 |

**c.** Fit a multi-group structural equation model (with country as the group variable) on the matrix of centered variables to investigate how the latent variables “social trust” and “trust in public institutions” affect the latent variable “well-being”. Estimate four versions of the multi-group structural equation model:

1) a configural measurement invariance model with country-specific regression coefficients in the regression equation of the structural model

2) a configural measurement invariance model with regression coefficients that are constrained to be equal across countries

3) a metric measurement invariance model with country-specific regression coefficients in the regression equation of the structural model

4) a metric measurement invariance model with regression coefficients that are constrained to be equal across countries

Compare the fit measures of the four estimated models and/or use model comparison tests to select the best model. Next discuss the results of this final model (e.g., model fit, estimated intercepts, (standardized) regression coefficients, etc.).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Chisq** | **Df** | **CFI** | **TLI** | **RMSEA** | **SRMR** | **AIC** |
| **config\_1** | 709.055 | 64 | 0.945 | 0.923 | 0.071 | 0.034 | 139051 |
| **config\_2** | 709.377 | 66 | 0.945 | 0.925 | 0.069 | 0.034 | 139047 |
| **metric\_1** | 775.804 | 74 | 0.94 | 0.927 | 0.068 | 0.043 | 139098 |
| **metric\_2** | 778.457 | 76 | 0.94 | 0.929 | 0.068 | 0.043 | 139096 |

> sem1<-'sotru =~NA\*+sotru1+sotru2+sotru3

+ truin =~NA\*truin1+truin2+truin3+truin4

+ webe =~NA\*webe1+webe2+webe3+webe4+webe5+webe6

+ sotru ~~1\*sotru

+ truin ~~1\*truin

+ webe ~~1\*webe

+ truin1 ~~ truin3

+ truin1 ~~ truin4

+ truin2 ~~ truin3

+ truin2 ~~ truin4

+ webe1 ~~ webe2

+ webe2 ~~ webe3

+ webe3 ~~ webe4

+ webe3 ~~ webe5

**d.**

> zess<- ess

> zess[,2:14]<-scale(ess[,2:14],center=TRUE,scale=FALSE)

>

> cancor.out<-cancor(cbind(fltdpr, fltsd, fltanx, wrhpp, enjlf, fltpcfl)

+ ~ppltrst+ pplfair+ pplhlp+ trstprl+ trstlgl+ trstplc+ trstplt, data=zess)

Warning message:

In model.matrix.default(mt, mf, contrasts) :

non-list contrasts argument ignored

> summary(cancor.out)

Canonical correlation analysis of:

7 X variables: ppltrst, pplfair, pplhlp, trstprl, trstlgl, trstplc, trstplt

with 6 Y variables: fltdpr, fltsd, fltanx, wrhpp, enjlf, fltpcfl

CanR CanRSQ Eigen percent cum scree

1 0.242629 5.887e-02 6.255e-02 77.37503 77.38 \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

2 0.110279 1.216e-02 1.231e-02 15.22875 92.60 \*\*\*\*\*\*

3 0.063206 3.995e-03 4.011e-03 4.96159 97.57 \*\*

4 0.041142 1.693e-03 1.696e-03 2.09741 99.66 \*

5 0.016167 2.614e-04 2.614e-04 0.32339 99.99

6 0.003343 1.118e-05 1.118e-05 0.01383 100.00

Test of H0: The canonical correlations in the

current row and all that follow are zero

CanR LR test stat approx F numDF denDF Pr(> F)

1 0.242629 0.92415 7.6396 42 18920 < 2.2e-16 \*\*\*

2 0.110279 0.98196 2.4539 30 16138 1.618e-05 \*\*\*

3 0.063206 0.99405 1.2056 20 13384 0.2378

4 0.041142 0.99804 0.6617 12 10678 0.7897

5 0.016167 0.99973 0.1834 6 8074 0.9815

6 0.003343 0.99999 NaN 2 NaN NaN

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Raw canonical coefficients

X variables:

Xcan1 Xcan2 Xcan3 Xcan4 Xcan5 Xcan6

ppltrst -0.108405 0.2853267 0.108076 -0.0495087 0.013165 -0.4336276

pplfair -0.282079 -0.4593998 -0.159970 0.1408960 -0.075323 -0.0624360

pplhlp -0.075600 0.2800853 -0.260620 0.1410840 0.205822 0.2527408

trstprl -0.097725 -0.1535614 0.320595 -0.0890089 0.367309 -0.0048426

trstlgl -0.080180 0.0573613 0.099537 -0.0092008 0.111209 0.2620818

trstplc -0.081954 -0.0060607 -0.126314 -0.4440381 -0.209336 -0.0123684

trstplt 0.067069 0.1206186 0.053729 0.3492442 -0.542852 0.0720115

Y variables:

Ycan1 Ycan2 Ycan3 Ycan4 Ycan5 Ycan6

fltdpr -0.38390 0.072259 0.26753 1.64697 -0.246829 0.463516

fltsd -0.20901 -0.016788 1.14942 -0.96373 0.813397 0.793988

fltanx -0.32340 1.188032 -0.46127 -0.10370 0.304109 -0.832190

wrhpp -0.22731 -0.035385 -1.35617 -0.33622 -0.046116 0.980694

enjlf -0.46398 -0.941008 0.22904 0.12495 0.474741 -1.137365

fltpcfl -0.18295 -0.034972 0.40065 -0.42358 -1.232723 -0.058797

>

> #redundancies

> redu<-redundancy(cancor.out)

> round(redu$Ycan,3)

Ycan1 Ycan2 Ycan3 Ycan4 Ycan5 Ycan6

0.030 0.002 0.000 0.000 0.000 0.000

>

> #computation redundancies from output

> R2tu<-cancor.out$cancor^2

> VAFYbyt<-apply(cancor.out$structure$Y.yscores^2,2,sum)/4

> redund<-R2tu\*VAFYbyt

> round(cbind(R2tu,VAFYbyt,redund,total=cumsum(redund)),3)

R2tu VAFYbyt redund total

Ycan1 0.059 0.770 0.045 0.045

Ycan2 0.012 0.192 0.002 0.048

Ycan3 0.004 0.132 0.001 0.048

Ycan4 0.002 0.131 0.000 0.048

Ycan5 0.000 0.154 0.000 0.048

Ycan6 0.000 0.122 0.000 0.048

**e.**

> samplesize<-dim(ess)[1]

> train<-ess[seq(2,samplesize,by=2),2:14]

> valid<-ess[seq(1,samplesize,by=2),2:14]

> train<-as.data.frame(scale(train,center=TRUE,scale=TRUE))

> valid<-as.data.frame(scale(valid,center=TRUE,scale=TRUE))

>

> #conduct CCA on training data

> cancor.train<-cancor(cbind(fltdpr, fltsd, fltanx, wrhpp, enjlf, fltpcfl)

+ ~ppltrst+ pplfair+ pplhlp+ trstprl+ trstlgl+ trstplc+ trstplt, data=train)

Warning message:

In model.matrix.default(mt, mf, contrasts) :

non-list contrasts argument ignored

>

> #conduct CCA on validation data

> cancor.valid<-cancor(cbind(fltdpr, fltsd, fltanx, wrhpp, enjlf, fltpcfl)

+ ~ppltrst+ pplfair+ pplhlp+ trstprl+ trstlgl+ trstplc+ trstplt, data=valid)

Warning message:

In model.matrix.default(mt, mf, contrasts) :

non-list contrasts argument ignored

>

> # canonical variates calibration set

> train.X1<-cancor.train$score$X

> train.Y1<-cancor.train$score$Y

>

> # compute canonical variates using data of calibration set and coefficients estimated on validation set

> train.X2<-as.matrix(train[,1:7])%\*%cancor.valid$coef$X

> train.Y2<-as.matrix(train[,8:13])%\*%cancor.valid$coef$Y

>

>

> #R(T,T\*) and R(U,U\*)

> round(cor(train.Y1,train.Y2)[1:3,1:3],3)

Ycan1 Ycan2 Ycan3

Ycan1 0.989 -0.111 -0.029

Ycan2 0.101 0.817 0.372

Ycan3 -0.053 -0.207 0.288

> round(cor(train.X1,train.X2)[1:3,1:3],3)

Xcan1 Xcan2 Xcan3

Xcan1 0.982 -0.042 -0.090

Xcan2 0.029 0.514 0.370

Xcan3 0.082 -0.095 0.405

>

> #R(U,T) and R(U\*,T\*)

> round(cor(train.X1,train.Y1)[1:3,1:3],3)

Ycan1 Ycan2 Ycan3

Xcan1 0.253 0.000 0.000

Xcan2 0.000 0.129 0.000

Xcan3 0.000 0.000 0.067

> round(cor(train.X2,train.Y2)[1:3,1:3],3)

Ycan1 Ycan2 Ycan3

Xcan1 0.246 -0.028 -0.001

Xcan2 -0.002 0.044 0.031

Xcan3 -0.019 0.039 0.043

>

> #R(T\*,T\*) and R(U\*,U\*)

> round(cor(train.Y2,train.Y2)[1:3,1:3],3)

Ycan1 Ycan2 Ycan3

Ycan1 1.000 -0.019 0.001

Ycan2 -0.019 1.000 0.015

Ycan3 0.001 0.015 1.000

> round(cor(train.X2,train.X2)[1:3,1:3],3)

Xcan1 Xcan2 Xcan3

Xcan1 1.000 -0.007 0.015

Xcan2 -0.007 1.000 -0.002

Xcan3 0.015 -0.002 1.000

**f.**

> as.matrix(round(cancor.out$structure$X.xscores[,1],3))

[,1]

ppltrst -0.676

pplfair -0.842

pplhlp -0.609

trstprl -0.544

trstlgl -0.599

trstplc -0.554

trstplt -0.441

> as.matrix(round(cancor.out$structure$Y.yscores[,1],3))

[,1]

fltdpr -0.740

fltsd -0.687

fltanx -0.677

wrhpp -0.749

enjlf -0.793

fltpcfl -0.644