GROUP 49



Luis Caminal Diaz, Ruifang Ge, Darragh Timothy O'Loughlin, Vittorio Balestrieri

Master in statistics and data sciencE, ku leuven

SOCIAL HAPPINESS AND LIFE SATISFACTION: A MULTIFACTOR STATISTICAL ANALYSIS

**TASK 1**

1. **Conduct confirmatory factor analysis on the covariance matrix of the items in Table 1 assuming a model with 5 correlated latent variables (one for each concept) and assuming each item only has loading on the concept it aims to measure. Discuss the results of the fitted CFA model (e.g., fit measures, (standardized) loadings, reliability of the factor scores, factor correlations, etc.). Next, conduct a multi-group analysis to assess whether metric and strong measurement invariance hold across the two countries and discuss the results.**

We fit a 5-factor model with correlated latent variables "Social Trust, Satisfaction Country, Job Autonomy, Job Satisfaction, Life Satisfaction". Below are the fit measures of the model.

Text

Description automatically generated with medium confidence

Our model is rejected by the absolute goodness of fit test (Chi-Squared=174.9, df= 44, p<0.000) However as the sample size is large this is not problematic since the high statistical power to detect a small deviation between the fitted model and the perfect model. Therefore, it is appropriate to rely on descriptive measures. The printed fit measures indicate that all measures meet the cut-off for a good fit, with CFI (0.985) and TLI (0.977) both greater than 0.95 and RMSEA (0.04) and SRMR (0.026) both less than 0.08.

We now look at the standardized solutions

A picture containing table

Description automatically generated

As shown in the standardized solution all variables have a significant positive correlation of at least .70 with the corresponding factor. This means that the individual variables have sufficient reliability, and that convergent validity is satisfied in the measurement model.

Furthermore, discriminant validity is also satisfied as the correlations between the latent factors are all significantly smaller than 1. Note that there is a rather strong correlation (.635) between the factors "Social Trust" and "Satisfaction Country". A high score between this factors makes sense because high life satisfaction seems difficult to obtain if there is low level of trust in the society in which you live and work. All other factors have weaker but positive correlations between each other.

From the output of the standardized solutions we can see that have reliablities respectively of 0.603 0.628 and 0.49. The Country satisfaction variables have ok reliabilities respectively 0.653, 0.755 and 0.63. The Job autonomy variables have reliabilities 0.812 and 0.551. The Job satisfaction variables have reliabilities 0.66 and 0.308 and the life satisfaction variables have reliabilities 0.861 and 0.607.

Now to fit multi-group models to investigate measurement invariance.

Text

Description automatically generated

The results of the analysis indicate that the fit of the configural model is similar to the fit obtained for the total data set. The configural model does meet the cutoff criteria of good fit, i.e. CFI and TLI are much above .95, and RMSE and MRSR are below .08.

We also see that imposing further model constraints (assume equal loadings across groups or assume equal loadings and intercepts across groups) results in a reduction of the fitness of the model but it still meets the cut-off criteria.

A LR test shows that metric measurement invariance is not supported by the data. In the same way, strong measurement invariance is not supported by the data

As the results of the analysis indicate that metric or strong measurement equivalence does not hold, we print the standardized loadings, the intercepts, and the factor correlations for the two countries from the standardized solution of the configural model.

A screenshot of a computer

Description automatically generated with medium confidence

Table

Description automatically generated

Both Countries have similar standardized loading implying that both countries have similar composite reliability.

All factor correlations are overall weak but positive in both Poland and Sweden.

*It could mean that we have too many factors involved in this model. To fix this problem I would suggest grouping Satisfaction of country and life together and grouping job Autonomy and Job satisfaction together.*

1. **Fit a structural equation model on the covariance matrix to investigate how the latent variables social trust, satisfaction country, job autonomy and job satisfaction affect the latent variable life satisfaction. Add also a structural relation between job autonomy and job satisfaction so that you can investigate both the direct effect of job autonomy on life satisfaction as well as the indirect effect of job autonomy on life satisfaction via job satisfaction. Discuss the results of the fitted structural equation model (e.g., model fit, (standardized) loadings, structural relations, etc.). Next, conduct a multi-group analysis to investigate whether the results of the structural equation model differ for the two countries. Test whether the population regression coefficients of the structural models (in the part “regressions”) are equal across the two countries.**

First, we fit the structural equation model and acquire the fit measures

Text

Description automatically generated with medium confidence

As before all our fit measures make the cut off for good fit and the model is rejected by absolute goodness of fit test as expected.

Next, we use the standardized loadings to compute the composite reliability of the factor scores for each country:

Text

Description automatically generated

Here we see Poland has some excellent composite reliabilities in Satisfaction Country, Job Autonomy, and Life Satisfaction whereas the composite reliabilities for Job Satisfaction and Social Trust have good but less strong results.

Sweden has 3 good composite reliabilities in Social Trust, Satisfaction Country, and Job Autonomy, 1 excellent CR in Life Satisfaction, and Job Satisfaction has a good CR.

Next, we inspect the standardized solution to see how the latent variables affect the life satisfaction of the country factors in both countries:

Table

Description automatically generated

The results of the structural relations indicate that in Poland, Life Satisfaction has a weak but significant positive correlation with each of the other 4 factors. Hence people who are satisfied in life have greater trust in society, job autonomy and are satisfied by their job and country.

Sweden also shows a weak but significant positive correlation between Life Satisfaction and the rest of the factors.

Now we conduct multi-group analysis to investigate whether the results of the structural equation model differ for the two countries.

Text

Description automatically generated

As could be expected, a LR test indicates that the assumption of having equal regression coefficients across the two countries is not supported by the data.

**Multi-group Analysis**

The results of the analysis indicate that the fit of the configural model is similar to the fit obtained for the total data set. The configural model meets the cut-off criteria of good fit, i.e. CFI and TLI are much above .95, and RMSE and MRSR are below .08.

We also see that imposing further model constraints (assume equal loadings across groups or assume equal loadings and intercepts across groups) results in a reduction of the fitness of the model but it still meets the cut-off criteria.

A LR test shows that metric measurement invariance is not supported by the data (LR=46.413, df=7, p<.001). In the same way, strong measurement invariance is not supported by the data (LR=109.62, df=14, p<.001).

As the results of the analysis indicate that metric or strong measurement equivalence do not hold, we print the standardized loadings, the intercepts, and the factor correlations for the two countries from the standardized solution of the configural model.

**TASK 2**

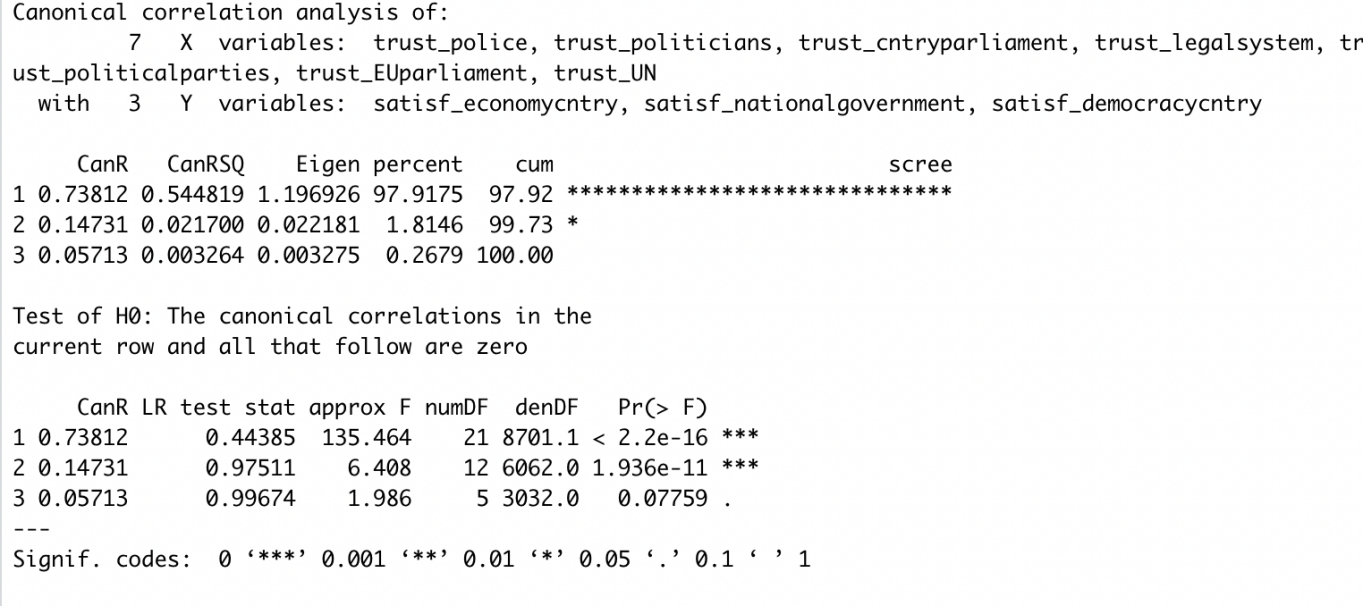
**Questions**

**a. Conduct a canonical correlation analysis on standardized variables to investigate the relations between the following two sets of variables:**

**Set of X variables: 7 items about trust in institutions.  
Set of Y variables: 3 items about satisfaction with the way the country is governed.**

**Discuss the results of the canonical correlation analysis. How many canonical correlations are significant? How much of the variance in the Y variables can be explained by the X variables? Which pairs of canonical variates are important? How can you interpret these canonical variates?**

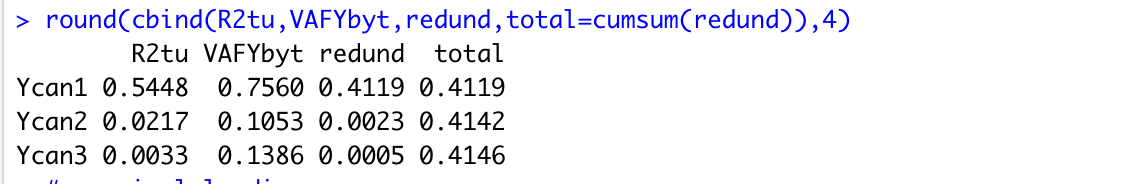
We standardize the variables and conduct the canonical correlation analysis. We print the results about significance of the canonical correlations and compute redundancies to see how much of the variance in the Y-variables is accounted for by the X-variables.



The output of the analysis shows that the last pair of canonical variates can be ignored since cannot be rejected.

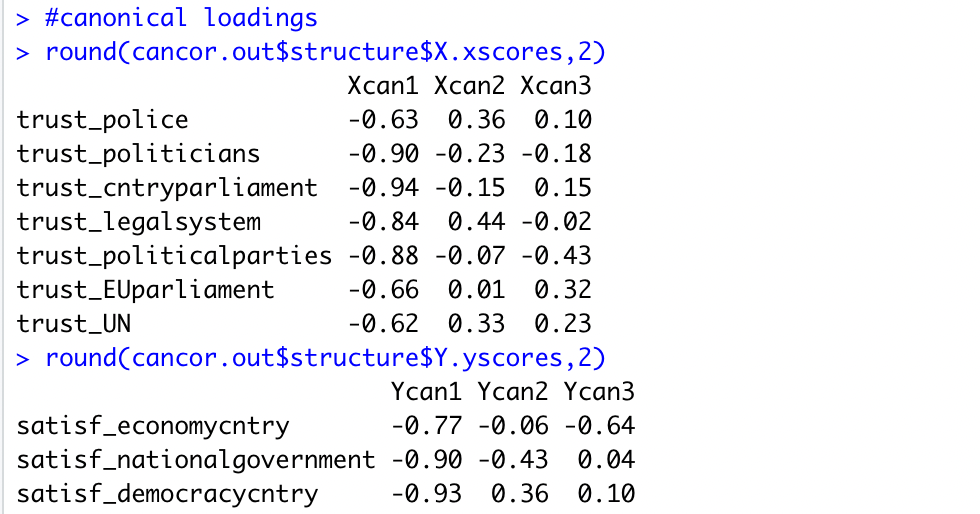
The first canonical correlation equals .738. This means that the canonical variate accounts for 54.48% of the variance in the canonical variate .

The second canonical correlation equals .147. The canonical variate accounts for 2.17% of the variance in the canonical variate.



Using the first two canonical variates (which are significant), we can see that the X variables account for 41.42% of the variance in the Y variables.

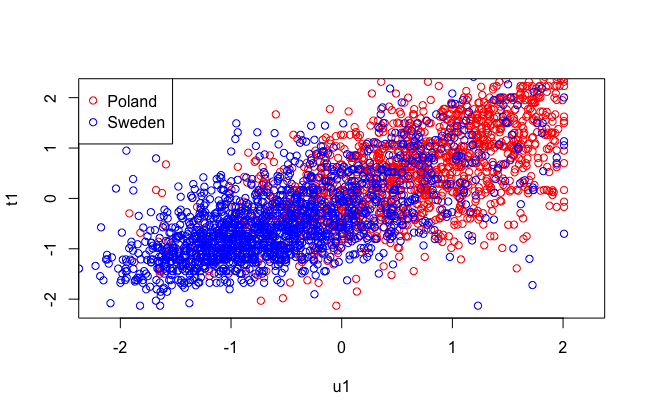
The first pair of canonical variates are practically relevant as accounts for 41.19% of the variance in the Y variables.



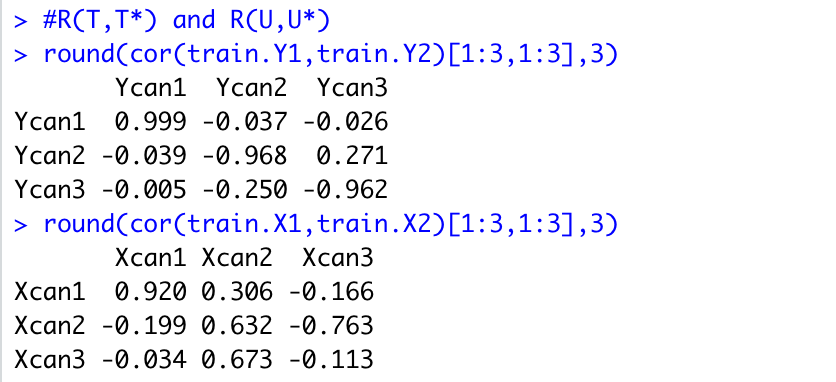
To interpret the first pair of canonical variates, we inspect the canonical loadings. As can be seen in the canonical loadings a higher score on means that the country’s parliament and politicians are less important, and a higher score on means that people are less satisfied with the country’s democracy and national government.

The canonical correlation analysis of people's trust in institutions and satisfaction with the country shows that the first canonical correlation coefficient is very significant, and the third canonical correlation coefficient is not significant. In the linear combination of satisfaction with the country, politics, parliament, law, political parties and people's satisfaction with the country show a high typical correlation coefficient.

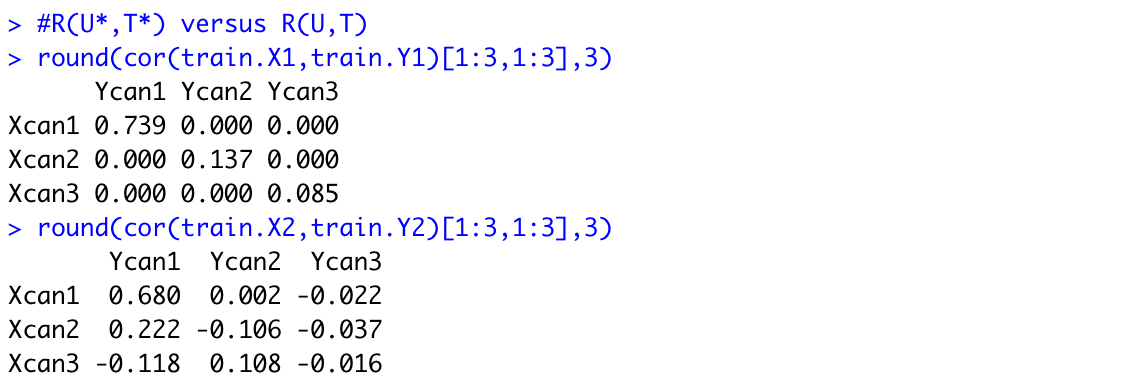
A plot of the first pair of canonical variates indicates that on average respondents of Poland find more satisfaction with the parliament and politicians whereas the opposite holds for Sweden. There is some overlap between the countries, and the top right corner of the graph contains some respondents from Sweden and has a high density of respondents from Poland.



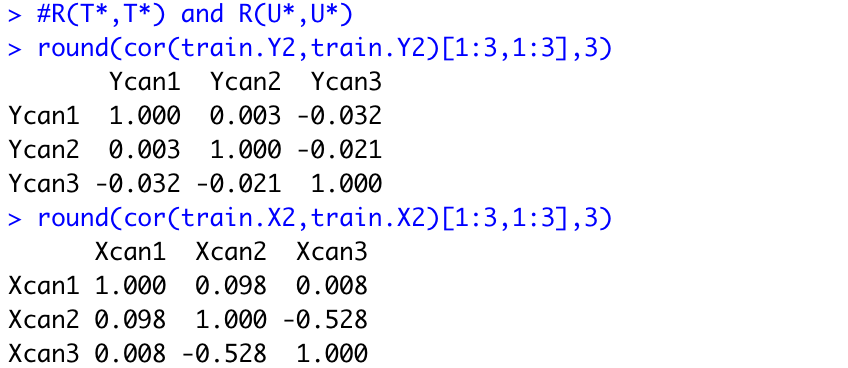
**b. Use the split-half approach to assess the validity of the solution. Assign even- numbered observations to the calibration set and assign odd-numbered observations to the validation set when conducting this analysis. Discuss what you can conclude about the validity of the solution.**



The first pair of canonical variates have excellent reliability and . However, the other pair of canonical variates do not have sufficient reliability. Off-diagonal elements in and are rather low and lower than diagonal elements. This is as expected because different canonical variates should be uncorrelated.



The comparison of and shows that the estimate of the first canonical correlation is rather stable (.74 is about equal to 0.68), whereas there is some capitalization on chance in the estimation of the other canonical correlations (.22 is larger than .11, .21 is larger than .058, etc).



Finally, canonical variates estimated on the validation data are more or less uncorrelated as off-diagonal elements of and are close to 0.

In sum, the redundancy analysis and the validation of the CCA using the split-half approach show that only the first pair of canonical variates are important and reliable. Hence, interpretation of the results should especially focus on the first pair of canonical variates.