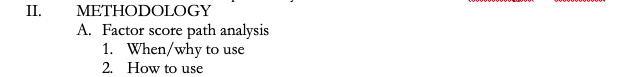
AAMAIO TERM PAPER

1. BACKGROUND
2. Croon’s Factor Score Regression
3. Explain (Croon, 2002)
4. FSR vs. SEM
5. Comparison of FSR methods vs. SEM (Devlieger, Mayer, & Rosseel, 2015)
6. Croon’s FSR + path analysis = alternative for SEM (Devlieger & Rosseel, 2017)
7. METHODOLOGY
8. Factor score path analysis
9. When/why to use
10. How to use
11. APPLICATION
12. Apply FSR to RA data (hopefully) & run path analysis
13. ALP data (8 factors)
14. MLQ data (7 factors)
15. Attempt SEM? **Maybe this is too far or irrelevant to the assignment?**
16. With the same data from FSR
17. RESULTS
18. Interpret results
19. Compare FSR to SEM?
20. Are they comparable?
21. CONCLUSION
22. Methodology’s relevance to I-O
23. How could it be used?
24. Overall impressions of the method
25. APPENDIX
26. R Code
27. Plots
28. Dataset
29. BACKGROUND

Over the past 100 years, structural equation modeling (SEM) has become one of the most popular and influential frameworks of quantitative methodologies in the social sciences (Kaplan, 2008). This is likely because of its ability to assign relationships between latent variables and observable variables, as this is essential to the field of psychometrics. Not only does SEM assess latent variables, via measurement models, it also simultaneously computes the causal relationships of these variables, via structural regression model (XX,XX). Ironically, this process is simultaneously the framework’s greatest strength and weakness. Misspecification at any point in an SEM, can result in bias that proliferates throughout the model, leading to unreliable estimates of parameters as well as regression coefficients (Hayes, 2019). Misinformed structural relationships can be detrimental for empirical research, as latent construct relationships are often the foundation of this practice (XX, XXXX). Hancock and Mueller argue that traditional methods are uncapable of detecting these misspecifications (2011), which leaves researchers looking for a more reliable analysis.

In the quest to eliminate the bias from their models, an increasing amount of researchers are switching from the simultaneous estimation method to multistage procedures (XX,XXXX). Different types of factor score regression (FSR) and factor score path analysis are among the stepwise methodologies growing in popularity (XX,XXXX). FSR is a two-step procedure in which each latent variable is measured separately in a structural regression model to create factor scores, then the factors scores are subsequently analyzed through a path analysis. It is crucial to be careful in removing bias from this data, as factor scores are not fully reliable (Grice, 2001). The two most common approaches are the *bias-avoiding approach* (Skrondal & Laake, 2001) and Croon’s *bias-correcting approach* (2002). Recent studies have found that Croon’s approach outperform Skrondal & Laake’s, due to its standardized coefficients and its inability to be extended into a path analytic framework (Devlieger et al., 2016; Lu et al., 2011). For these reasons, this paper will focus on Croon’s bias-correcting FSR approach.

Croon’s approach to factor score regression, is based on the premise that there is a difference between the variances and covariances of the factor scores and the variances and covariances of the true latent scores (2002). The approach uses an estimation of the variances and covariances of the true latent scores, rather than factor scores, in order to estimate the parameters (Croon, 2002). Between the studies of Croon (2002), Lu et al. (2011), and Devlieger et al. (2016), the approach has shown to be effective in producing unbiased parameter estimates in both populations and finite samples. When the sample size is large, the method was shown to be comparable in efficient, mean square error, power, and type-1 error rate as SEM (Devlieger et al., 2016). In 2017, Devlieger et al. found that Croon’s method performs just as well as SEM with regard to bias and convergence rate when path analysis is used. They also found evidence that Croon’s method handles misspecifications better than SEM and requires a smaller sample size (Devlieger et al., 2017). Hayes and Usami (2019) found that Croon’s method outperformed SEM when it is using standard specification of unique factor covariances. SEM performed comparably well when the unique factor covariances were specified, but was outperformed again when the SEM specified the unique factor covariances, but misspecified the structural model (Hayes, 2019). Evidence provided by the aforementioned researchers gives reason to believe that Croon’s FSR approach may be a suitable alternative for SEM in certain instances (e.g., misspecifications in parameter or structure, small sample size). In this paper I will further describe Croon’s FSR approach, utilize its methodology using real-life data, and assess its applicability in the field of Industrial-Organization psychology.



Croon’s Method

Croon’s bias correcting factor score regression method was developed by Marcel Croon (2002) to combat the biases that are inherent in factor scores. It is based off of the idea that there is a difference between the variances and covariances of the factor scores and the true latent variable scores. In order to eliminate this bias, Croon (2002) uses an estimation of the true latent variable scores, rather than the factor scores, when estimating the regression parameters (Devlieger & Rosseel, 2017). The stepwise approach is as follows: (1) Factor scores are first computed using the regression predictor (Thomson, 1934; Thurstone, 1935) or the Bartlett predictor (Bartlett, 1937; Thomson, 1938); (2) The variances and covariances of theses scores are calculated, which are then used to; (3) compute the variances and covariances of the true latent variables; and (4) these estimates are then used to calculate the regression coefficient. Because the variances and covariances are unbiased, so is the regression coefficient estimate (Devlieger et al., 2015). Using the unbiased variance and covariance estimates, a path analysis is performed using the estimated variances and covariances. Devlieger and Rosseel (2017) combined Croon’s method with path analysis, which allows non-recursive models (i.e., models that contain feedback loops or reciprocal effects) to be analyzed using factor scores.

In order to use Croon’s method for hypothesis testing, it is necessary to have a corresponding significance test, which requires a standard error and a theoretical distribution. Devlieger and colleagues (2015) developed a method for calculating the standard error of Croon’s approach that corresponds with the corrected regression coefficient. This adjusted standard error is created by calculating the prediction error in the factor scores. The adjusted standard error can be for this approach can be calculated as:

The newest addition to factor score regression is a set of fit indices and a model comparison test. In order to inspect model fit, Devlieger et al. (2019) propose fit indices for factor score regression based on the X^2, RMSEA, SRMR, and CFI fit indices used for structural equation modeling. Their newly proposed X^2 can be used to conduct a model comparison test, which not only allows us to see how well the model fits Croon’s method, but how it holds up to other approaches (e.g., SEM).

Limitations

As this bias-correcting method continues to be explored, researchers have uncovered limitations to the approach that should be researched more in depth. For example, factor score regression method only works when there are at least three items per latent variable. As of now, the Croon approach does not work for connected measurement models (e.g., models with cross-loadings or correlated residual errors). Devlieger and Rosseel (2017) are currently working to enhance the method’s capability of handling these types of models as well as extend the inferences made by the model. Because of the method’s age, there is very little software available to perform it – making it a difficult and tedious process. Many of the functions used for this method are still being developed in R by the aforementioned authors, and many functions are not yet available (e.g., model fit indices).

1. APPLICATION
2. Apply FSR to RA data (hopefully) & run path analysis
3. ALP data (8 factors)
4. MLQ data (7 factors)

For ease of application, this paper will test Croon’s method on a simple regression model of latent variables. Particularly, we will look at the relationship between two latent variables, Idealized Attributes (IA; Bass & Avolio, 20XX) and Purpose (Apter, XXXX), and their observed variables (i.e., the items used to measure the latent constructs).

Idealized Attributes is a factor of Transformational Leadership, as measured by the Multifactor Leadership Questionnaire (MLQ; Bass and Avolio). It subcomponent of the Idealized Influence construct. Idealized Influence describes leaders that act as role models and are well trusted by their followers. They often show a strong sense of morality and set high ethical standards. The component of Idealized Attributes differs from its counterpart, *Idealized Behaviors*, in that it refers to the attributes that a leader allows their followers to observe (Northouse, 2016). IA consists of 4 items on the MLQ:

1. Item 1
2. Item 2
3. Item 3
4. Item 4

Purpose is a microclimate corresponds with reversal theory’s Means-Ends’ “Telic” motivational state, which encourages the pursuit and achievement of organizational and individual goals as its contribution (Carter, 2009). Leaders that foster a strong microclimate of Purpose encourage their followers to focus on the goals and see things in the long-term. Achievement is valued and there is a collective sense of vision and strategy. Purpose consists of 5 items on the Apter Leadership Profiling System:

1. Awareness of the team’s strategic purpose
2. Inspired commitment to long-term goals
3. A sense of mission
4. Anticipation of future consequences
5. A clear vision for the team

I expect these two factors to converge highly, as they appear to measure somewhat similar latent constructs. Second, I want to see if purpose predicts IA. As for the manifest variables (i.e., the items used to measure the latent constructs), I expect \_\_\_\_\_\_\_\_.

Analysis

Step 1: Compute factor scores using regression or Bartlett predictor

Step 2: Calculate variance and covariance

Step 3: Estimate variance and covariance of the true latent variability scores

Step 4: Perform path analysis

The dataset was analyzed using Roosseel’s Lavaan package in R. I created my hypothesized model by using the Lavaan model language:

model <-  
 '# latent variable definitions  
  
 purpose =~ P1 + P2 + P3 + P4 + P5  
 IA =~ IA1 + IA2 + IA3 + IA4

# regressions

purpose ~ IA '

Our model consists of “purpose” and its corresponding variables, as well as “IA” and its corresponding variables. The “model” object was used for the next steps of the analysis. The function “lavaan::: sam” automates the four steps of Croon’s factor score regression method. To utilize this function, I input my model object. After 14 iterations, my model ended normally. The

> summary(fit.sam, standardized = TRUE, fit.measures = TRUE)  
lavaan 0.6-7 ended normally after 14 iterations  
  
Parameter Estimates:  
  
 Standard errors Twostep  
 Information Expected  
 Information saturated (h1) model Structured  
  
Regressions:  
 Estimate Std.Err z-value P(>|z|) Std.lv Std.all  
 purpose ~   
 IA 1.742 0.724 2.407 0.016 1.742 0.689  
  
Variances:  
 Estimate Std.Err z-value P(>|z|) Std.lv Std.all  
 .purpose 0.277 0.093 2.971 0.003 0.277 0.525  
 IA 0.083 0.061 1.357 0.175 0.083 1.000

To test the fit of the model, I used the formula from Devlieger (2019).

Results

I used Croon’s bias-reducing factor score regression method to test the relationship between two latent variables (Idealized Attributes and Purpose).

The data for this example included scores of leaders on the MLQ and ALP. To measure Idealized Attributes, I used items IA1, IA2, IA3, and IA4. To measure Purpose, I used items P1, P2, P3, P4, and P5. Both tests were scored on a Likert scale from 1 to 5. Descriptives for the observed variables are found in Table 1.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Mean** | **Standard Deviation** |  | **Mean** | **Standard Deviation** |
| IA1 | 2.64 | 0.83 | P1 | 5.10 | 0.85 |
| IA2 | 3.44 | 0.61 | P2 | 4.95 | 0.96 |
| IA3 | 3.17 | 0.57 | P3 | 4.90 | 0.98 |
| IA4 | 2.74 | 0.89 | P4 | 5.27 | 0.80 |
|  |  |  | P5 | 4.85 | 1.03 |

As I assumed, purpose predicted IA (r = 0.69)

FSR vs. SEM

After running the structural equation model for my

Appendix

model <-  
 '# latent variable definitions  
  
 purpose =~ P1 + P2 + P3 + P4 + P5  
 IA =~ IA1 + IA2 + IA3 + IA4

# regressions

purpose ~ IA '

> summary(fit.sam, standardized = TRUE, fit.measures = TRUE)  
lavaan 0.6-7 ended normally after 14 iterations  
  
Parameter Estimates:  
  
 Standard errors Twostep  
 Information Expected  
 Information saturated (h1) model Structured  
  
Regressions:  
 Estimate Std.Err z-value P(>|z|) Std.lv Std.all  
 purpose ~   
 IA 1.742 0.724 2.407 0.016 1.742 0.689  
  
Variances:  
 Estimate Std.Err z-value P(>|z|) Std.lv Std.all  
 .purpose 0.277 0.093 2.971 0.003 0.277 0.525  
 IA 0.083 0.061 1.357 0.175 0.083 1.000

References

Kaplan, D. (2008). Structural Equation Modeling: Foundations and Extensions (2nd ed.). SAGE. ISBN 978-1412916240.

Hayes, T., & Usami, S. (2020). Factor score regression in the presence of correlated unique factors. *Educational and Psychological Measurement*, *80*(1), 5-40.

Hancock, G. R., & Mueller, R. O. (2011). The reliability paradox in assessing structural relations within covariance structure models. *Educational and Psychological Measurement*, *71*(2), 306-324.