A comparison of data analyses with WarpPLS and Stata: A study of trust and its role regarding internet use and subjective well-being

Ayman Samak
Texas A&M International University, USA

Muhammad Rofiqul Islam
Texas A&M International University, USA

Doreen Hanke
McNeese State University, USA

Abstract

This study investigates the mediating roles of social and institutional trust in the relationship between internet use and subjective well-being, using partial least squares (PLS)-based structural equation modeling (SEM). We compare WarpPLS 8.0 and Stata's PLS-SEM package, utilizing data from the European Social Survey (ESS), round 8. Our results show consistent model fit and path coefficients across both tools, confirming the significant mediating effects of trust. WarpPLS stands out for its advanced model diagnostics, while Stata's PLS-SEM excels in integrating with Stata's broader data management and statistical analysis tools. This comparative analysis contributes to the SEM methodological literature.

Keywords: Comparative Analysis; Structural Equation Modeling; Partial Least Squares; PLS-SEM; Stata; WarpPLS.

Introduction

Structural equation modeling (SEM) is a pivotal statistical methodology that allows researchers to assess hypothesized relationships among observed and latent variables (Bollen, 2014; Kline, 2023; Tarka, 2018). Partial least squares SEM (PLS-SEM) is particularly suited for exploratory research. The choice of software profoundly influences analytical outcomes, especially when handling mediation effects and multi-country data. This study compares two prominent SEM tools: WarpPLS and Stata's package for PLS-SEM.

WarpPLS 8.0 (Kock, 2023) is an advanced SEM tool recognized for implementing traditional composite-based and contemporary factor-based PLS-SEM algorithms. It is designed to handle non-linear relationships and calculate stable p-values, making it particularly valuable for complex SEM analyses (Amora, 2023; Hubona & Belkhamza, 2021; Kock, 2015, 2020, 2023).

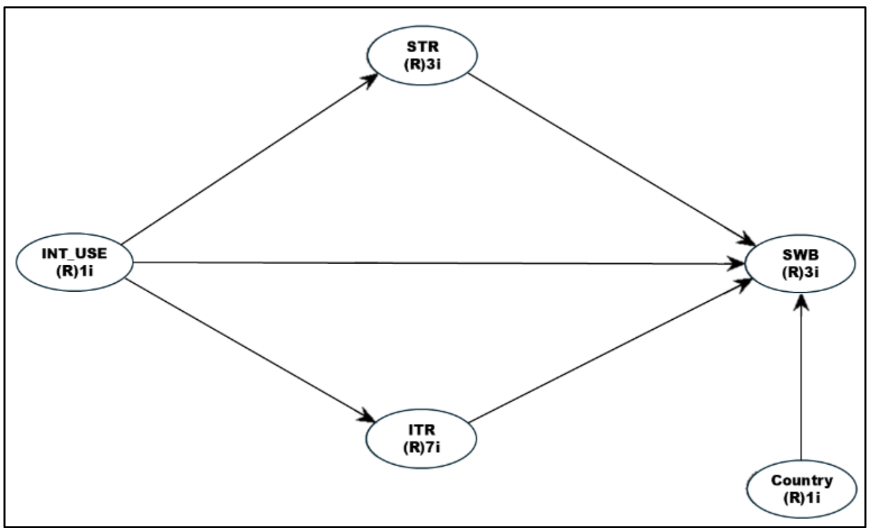
Its advanced features, such as its support for nonlinear SEM and robust model fit indices, enhance its adaptability for various research designs (Kock, 2023).

Stata's PLS-SEM package offers integration with Stata's comprehensive statistical, graphical, and data management tools, providing a seamless workflow for users accustomed to Stata's environment (Mehmetoglu & Venturini, 2021). It is adaptable, offering support for large datasets and complex models, which makes the package suitable for rigorous empirical research (Venturini & Mehmetoglu, 2019). Furthermore, continuous updates and support ensure that researchers have access to the latest methodological advancements.

The choice between both tools is not trivial, as it hinges on their ability to provide accurate, reliable, and interpretable models. This comparative analysis will utilize empirical data to highlight each software's results, contributing to the broader methodological literature on SEM.

Illustrative model and data

Figure 1: Illustrative model used



Notes: INT_USE = Internet Use; ITR = Institutional trust; STR = Social trust; SWB = Subjective well-being; notation under latent variable acronym describes measurement approach and number of indicators, e.g., (R)3i = reflective measurement with three indicators.

The study employed ESS - round 8[1] data, disseminated between 2016 and 2018. Before analysis, the dataset went through a cleaning process following Smith and Noble (2014). The final sample size was 15,238 data points, representing respondents from 21 European countries[2]. For the analytical approach, we adopted PLS-SEM, executed through the WarpPLS 8.0 software

[1] A rigorous multi-country survey conducted through face-to-face interviews since 2001, the ESS provides data on attitudes, beliefs, and behaviors across Europe (ESS ERIC, 2023). For more information, please visit https://www.europeansocialsurvey.org/.

[2] The countries include Austria, Belgium, Switzerland, the Czech Republic, Germany, Estonia, Spain, Finland, France, the United Kingdom, Hungary, Ireland, Iceland, Italy, Lithuania, the Netherlands, Norway, Poland, Portugal, Sweden, and Slovenia.

(Kock, 2023) and Stata's PLS-SEM (Mehmetoglu & Venturini, 2021). The model depicted in Figure 1 reflects the hypothesized relationships among variables.

Results

Tables 1, 2, and 3 show the data analysis results from WarpPLS and Stata's PLS-SEM. Figures 2 and 3 show comparable structural models of the WarpPLS output and the modified Stata output, respectively. The original Stata output displays a simple model of causal relationships between variables. It lacks additional diagnostics, such as path coefficients, p-values, or R-squared details, making it less informative than WarpPLS's output.

Table 1: Indicator loadings and cross-loadings of items

	ITR		STR		SWB	
	WarpPLS	STATA	WarpPLS	STATA	WarpPLS	STATA
ITR1	**(0.799)**	**(0.815)**	0.402	0.402	0.278	0.275
ITR2	**(0.656)**	**(0.679)**	0.338	0.338	0.263	0.257
ITR3	**(0.873)**	**(0.859)**	0.394	0.394	0.245	0.242
ITR4	**(0.797)**	**(0.789)**	0.319	0.320	0.225	0.225
ITR5	**(0.754)**	**(0.752)**	0.328	0.329	0.221	0.220
ITR6	**(0.865)**	**(0.851)**	0.385	0.385	0.247	0.243
ITR7	**(0.855)**	**(0.857)**	0.403	0.404	0.282	0.279
STR1	0.407	0.409	**(0.842)**	**(0.856)**	0.287	0.286
STR2	0.376	0.379	**(0.847)**	**(0.855)**	0.290	0.287
STR3	0.357	0.358	**(0.803)**	**(0.778)**	0.246	0.241
SWB1	0.306	0.309	0.320	0.320	**(0.887)**	**(0.887)**
SWB2	0.269	0.272	0.292	0.293	**(0.885)**	**(0.885)**
SWB3	0.176	0.177	0.183	0.185	**(0.657)**	**(0.657)**

Notes: ITR = Institutional trust; STR = Social trust; SWB = Subjective well-being; WarpPLS = Structure loadings and cross-loadings; STATA = Outer loadings and Cross-loadings

Table 2: Correlation between latent variables and errors

	ITR		STR		SWB	
	WarpPLS	STATA	WarpPLS	STATA	WarpPLS	STATA
ITR	(0.803)	(0.802)	0.458	0.460	0.312	0.313
STR	0.458	0.460	(0.831)	(0.830)	0.331	0.328
SWB	0.312	0.313	0.331	0.328	(0.817)	(0.817)

Notes: ITR = Institutional trust; STR = Social trust; SWB = Subjective well-being; WarpPLS = correlation between latent variables vs. square roots of AVEs; STATA = Squared interfactor correlation vs. AVEs

Table 3: Latent variable coefficients

	ITR		STR		SWB	
	WarpPLS	STATA	WarpPLS	STATA	WarpPLS	STATA
Composite reliability	0.927	0.926	0.870	0.776	0.855	0.742
Cronbach's alpha	0.907	0.907	0.776	0.776	0.742	0.742
AVE	0.645	0.644	0.691	0.690	0.667	0.665
Full Collin. VIFs	1.371		1.367		1.256	

Notes: ITR = Institutional trust; STR = Social trust; SWB = Subjective well-being

Figure 2: WarpPLS Output

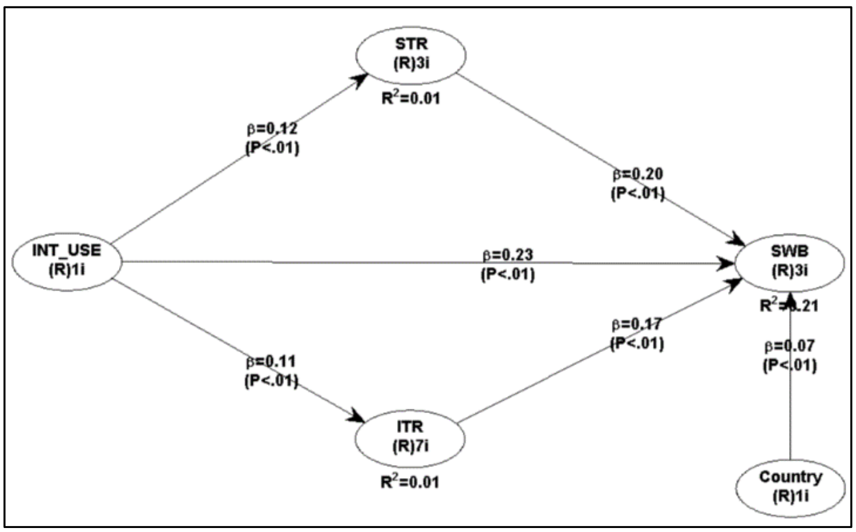
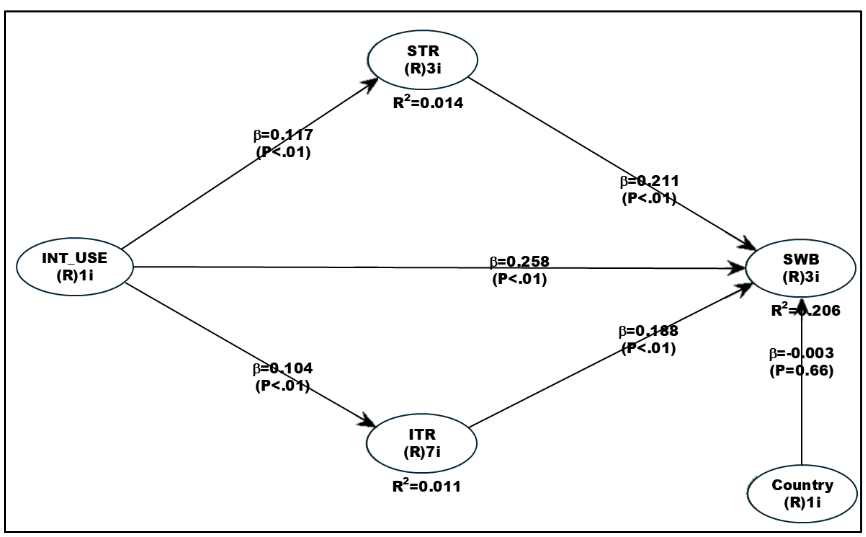


Figure 3: Stata output (modified)



WarpPLS and Stata's PLS-SEM produced consistent results in model fit, path coefficients, and validation indices. WarpPLS provided detailed model fit and quality indices, such as

Tenenhaus GoF and collinearity VIFs. It also offers greater flexibility in output selection, allowing choices like combined, pattern, and structure loadings and correlation evaluations with VIFs. At the same time, Stata is limited to a single output with outer loadings and interfactor correlations. Stata's PLS-SEM provided detailed bootstrapping results. However, the reliance on bootstrapping in PLS-SEM can sometimes lead to instability in p-values, particularly when working with large datasets (Kock, 2018). In contrast, WarpPLS addresses this issue using more stable p-value calculation methods.

Conclusion

WarpPLS and Stata's PLS-SEM demonstrate their effectiveness in analyzing the mediating roles of social and institutional trust between internet use and subjective well-being. WarpPLS offers distinct advantages in advanced diagnostics, stable p-value calculations, and handling nonlinearity, making it ideal for more complex SEM models. In contrast, Stata's PLS-SEM provides flexibility through its integration with broader statistical tools, catering to users needing multifaceted analytical techniques. The choice between these tools hinges on factors such as the complexity of model diagnostics, desired output, dataset size, and the type of analysis being conducted.

Acknowledgments

The authors are thankful to the editor of this journal, Dr. Ned Kock, for his invaluable guidance and dedication to the continued development of WarpPLS and his valuable feedback for this article.

References

Amora, J. T. (2023). On the validity assessment of formative measurement models in PLS-SEM. *Data Analysis Perspectives Journal*, *4*(2), 1–7.

Bollen, K. A. (2014). *Structural equations with latent variables*. New York, NY: John Wiley & Sons.

ESS ERIC. (2023). *ESS8—Integrated file, edition 2.3* [Dataset]. Sikt - Norwegian Agency for Shared Services in Education and Research. https://doi.org/10.21338/ess8e02_3

Hubona, G., & Belkhamza, Z. (2021). Testing a moderated mediation in PLS-SEM: A full latent growth approach. *Data Analysis Perspectives Journal*, *2*(4), 1–5.

Kline, R. B. (2023). *Principles and practice of structural equation modeling*. New York, NY: Guilford Publications.

Kock, N. (2015). One-tailed or two-tailed P values in PLS-SEM? *International Journal of E-Collaboration*, *11*(2), 1–7.

Kock, N. (2018). Should bootstrapping be used in pls-sem? Toward stable p-value calculation methods. *Journal of Applied Structural Equation Modeling*, *2*(1), 1–12.

Kock, N. (2020). Full latent growth and its use in PLS-SEM: Testing moderating relationships. *Data Analysis Perspectives Journal*, *1*(1), 1–5.

Kock, N. (2023). *WarpPLS User Manual: Version 8.0*. Laredo, TX: ScriptWarp Systems.

Mehmetoglu, M., & Venturini, S. (2021). *Structural equation modelling with partial least squares using Stata and R*. London, England: Chapman and Hall/CRC.

Smith, J., & Noble, H. (2014). Bias in research. *Evidence-Based Nursing*, *17*(4), 100–101.

Tarka, P. (2018). An overview of structural equation modeling: Its beginnings, historical development, usefulness and controversies in the social sciences. *Quality & Quantity*, *52*(1), 313–354.

Venturini, S., & Mehmetoglu, M. (2019). Plssem: A Stata package for structural equation modeling with partial least squares. *Journal of Statistical Software*, *88*(1), 1–35.