Book of Short Papers SIS 2020

Editors: Alessio Pollice, Nicola Salvati and Francesco Schirripa Spagnolo

Published by Pearson

www.pearson.com

ISBN 9788891910776

Longitudinal data analysis using PLS-PM approach

Analisi dei dati longitudinali attraverso l'approccio PLS-PM

Rosanna Cataldo, Corrado Crocetta, Maria Gabriella Grassia, Marina Marino

Abstract Longitudinal data over the past 20 years have seen a greater diffusion in the social sciences. Accompanying this growth was an interest in the methods for analyzing such data. Structural Equation Modeling (SEMs) and especially Partial Least Squares Path Modeling (PLS-PM) are a valuable way to analyze longitudinal data because it is both flexible and useful for answering common research questions. The aim of this paper is to demonstrate how PLS-PM can help us to analyze longitudinal data.

Abstract *I dati longitudinali degli ultimi 20 anni hanno visto una maggiore diffusione nelle scienze sociali. Ad accompagnare questa crescita è stato un interesse per i metodi di analisi di tali dati. I modelli ad equazioni strutturali (SEM) ed in particolare il metodo Partial Least Squares- Path Modeling (PLS-PM) sono un prezioso metodo di analizzare i dati longitudinali perché è sia flessibile che utile per rispondere a domande di ricerca comuni. Lo scopo di questa ricerca è di dimostrare come questo approccio può aiutarci ad analizzare i dati longitudinali.*

Key words: Partial Least Squares - Path Modeling, Longitudinal study

---

Rosanna Cataldo
Department of Social Science, University of Naples Federico II, vico Monte della Pietà, 1, 80138 Napoli, Italy e-mail: rosanna.cataldo2@unina.it
Corrado Crocetta
Department of Economics, University of Foggia, Largo Papa Giovanni Paolo II, 71121 Foggia, Italy e-mail: corrado.crocetta@unifg.it
Maria Gabriella Grassia
Department of Social Science, University of Naples Federico II, vico Monte della Pietà, 1, 80138 Napoli, Italy e-mail: mgrassia@unina.it
Marina Marino
Department of Social Science, University of Naples Federico II, vico Monte della Pietà, 1, 80138 Napoli, Italy e-mail: mari@unina.it

1 Introduction

A longitudinal study refers to an investigation where participant outcomes are collected at multiple times. Longitudinal data are difficult to collect, but longitudinal research is popular. And this popularity seems to be growing. With this comes the subsequent need for good data analysis methods to analyze these special kinds of data. The analysis of change in longitudinal data has attracted considerable attention over past decades in behavioral research [1]; [3]. Researchers have proposed a number of advanced and sophisticated quantitative approaches to address the stability and change of variables over time [5]; [17]. A broad range of statistical methods exists for analyzing data from longitudinal designs. Each of these methods has specific features and the use of a particular method in a specific situation depends on such things as the type of research, the research question, and so on. The central concern of longitudinal research, however, revolves around the description of patterns of stability and change, and the explanation of how and why change does or does not take place [11]. A common design for longitudinal research in the social sciences is the panel or repeated measures design, in which a sample of subjects is observed at more than one point in time. If all individuals provide measurements at the same set of occasions, we have a fixed occasions design. When occasions are varying, we have a set of measures taken at different points in time for different individuals. Such data occur, for instance, in growth studies, where individual measurements are collected for a sample of individuals at different occasions in their development. The model on longitudinal data can be approached from several perspectives, and the model can be constructed as a Structural Equation Model (SEM). According to Baltes and Nesselroade [2], SEM is a valuable way to analyze longitudinal data because it is both flexible and usefulfor answering common research questions. The explicit invocation of latent variables (LVs) afforded by the SEM makes this framework the one most commonly used to implement and analyze longitudinal data [18]; [23]. The SEM framework is often used to study change processes because this framework provides an opportunity to specify multiple latent variables as predictors and outcomes. SEM techniques include two main methods: covariance-based SEM (CB-SEM), represented by LISREL [14] and component-based SEM, with Partial Least Square-Path Modeling (PLS-PM) [9]; [24]. PLS-PM can be used to implement and to analyze the longitudinal data.

2 Longitudinal data analysis with PLS-PM approach

Recently, Roemer [20] has proposed using the component-based approach to SEM - PLS-PM in a longitudinal study [7]; [24]. In accordance with Roemer [20], we posit that PLS path modeling is highly appropriate for an analysis of the development and change in constructs in longitudinal studies, since it offers three favorable methodological characteristics. First, constructs often need to be predicted in evolutionary models [12]; [22]. Secondly, model complexity quickly increases when development and change need to be analyzed in longitudinal studies. This is due to the larger number of constructs that are measured at different points in time and the respective effects between those constructs [12]. PLS-PM is well suited to dealing with such complex models [8]; [28]. Thirdly, sample sizes can become quite small in longitudinal studies [13]. PLS path modeling is particularly appropriate in such cases [10]; [15]. In Wold's original design of PLS-PM [27] it was expected that each construct would necessarily be connected to a set of observed variables. On this basis, Lohmöller [16] proposed a procedure to treat hierarchical constructs, the so-called hierarchical component model. There are several main reasons for the inclusion of a Higher-Order Construct Model: Higher-Order Construct Models prove valuable if the constructs are highly correlated; the estimations of the structural model relationships may be biased as a result of collinearity issues, and a discriminant validity may not be established. In situations characterized by collinearity among constructs, a Second-Order Construct can reduce such collinearity issues and may solve discriminant validity problems. PLS path modeling allows for the conceptualization of a hierarchical model, through the use of three main approaches existing in the literature: the Repeated Indicators Approach [16], the Two Step Approach [19] and the Mixed Two Step Approach [4]. The Repeated Indicators Approach is the most popular approach when estimating Higher-Order Constructs in PLS-PM [25]. We propose using a Higher Order Model to analyze longitudinal data. An example, with three points in time, is presented in Fig. 1.

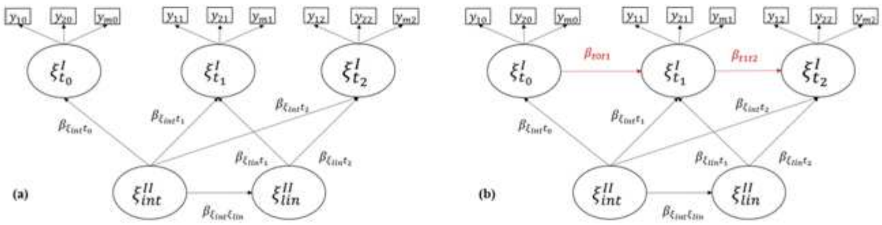

Fig. 1 (a) PLS-PM model - three times with m indicators in each n times; (b) PLS-PM model - three times with m indicators in each n times, with $\xi^I_{t_{i-1}}$ that impact on $\xi^I_{t_i}$

The Higher Order LV ξ^{II}_{int} describes the mean growth, and the LV ξ^{II}_{lin}.the mean slope. ξ^{II}_{int} is reflected in the construct of first order $\xi^I_{t_0}$, $\xi^I_{t_1}$....$\xi^I_{t_n}$. The construct of second order ξ^{II}_{lin} is reflected in the construct of first order $\xi^I_{t_1}$....$\xi^I_{t_n}$, The equations of the inner model are:

$$
\begin{aligned}
&\xi_{lin}^{II} = \beta_{0lin} + \beta_{\xi_{int}\xi_{lin}}\xi_{int}^{II} + \zeta_{lin}\\
&\xi_{t_0}^{I} = \beta_{0t_0} + \beta_{\xi_{int}t_0}\xi_{int}^{II} + \zeta_{t_0}\\
&\xi_{t_1}^{I} = \beta_{0t_1} + \beta_{\xi_{int}t_1}\xi_{int}^{II} + \beta_{\xi_{lin}t_1}\xi_{lin}^{II} + \zeta_{t_1}\\
&\dots\\
&\xi_{t_i}^{I} = \beta_{0t_i} + \beta_{\xi_{int}t_i}\xi_{int}^{II} + \beta_{\xi_{lin}t_i}\xi_{lin}^{II} + \zeta_{t_i}\\
&\xi_{t_n}^{I} = \beta_{0t_n} + \beta_{\xi_{int}t_n}\xi_{int}^{II} + \beta_{\xi_{lin}t_1}\xi_{lin}^{II} + \zeta_{t_n}
\end{aligned}
\tag{1}
$$

where $\beta_{\xi_{int}\xi_{lin}}$ is the strength and sign of the relations between construct ξ_{lin}^{II} and the predictor construct ξ_{int}^{II}; $\beta_{\xi_{int}\xi_{lin}}$ representing the growth mean rate; $\beta_{\xi_{int}t_i}$ is the strength and sign of the relations between construct $\xi_{t_i}^{I}$ and the predictor construct ξ_{int}^{II}; $\beta_{\xi_{lin}t_i}$ is the strength and sign of the relations between construct $\xi_{t_i}^{I}$ and the construct ξ_{lin}^{II}. They indicate how both intercept and slope factors contribute to explaining each time. β_0 is just the intercept term and ζ accounts for the residuals. The intercept term β_0 of each equation should always be non-significant. If we introduce the impact of the LV at i-1 time ($\xi_{t_{i-1}}^{I}$) on the LV at i time ($\xi_{t_i}^{I}$), for its better prediction ($\xi_{t_0}^{I} \rightarrow \xi_{t_1}^{I}$; $\xi_{t_1}^{I} \rightarrow \xi_{t_2}^{I}$;....; $\xi_{t_{n-1}}^{I} \rightarrow \xi_{t_n}^{I}$) as in Fig. 1 (b), the equations of the inner model become:

$$
\begin{aligned}
&\xi_{lin}^{II} = \beta_{0lin} + \beta_{\xi_{int}\xi_{lin}}\xi_{int}^{II} + \zeta_{lin}\\
&\xi_{t_0}^{I} = \beta_{0t_0} + \beta_{\xi_{int}t_0}\xi_{int}^{II} + \zeta_{t_0}\\
&\xi_{t_1}^{I} = \beta_{0t_1} + \beta_{\xi_{int}t_1}\xi_{int}^{II} + \beta_{\xi_{lin}t_1}\xi_{lin}^{II} + \beta_{t0t1}\xi_{t_0}^{I} + \zeta_{t_1}\\
&\dots\\
&\xi_{t_i}^{I} = \beta_{0t_i} + \beta_{\xi_{int}t_i}\xi_{int}^{II} + \beta_{\xi_{lin}t_i}\xi_{lin}^{II} + \beta_{t_{i-1}t_i}\xi_{t_{i-1}}^{I} + \zeta_{t_i}\\
&\xi_{t_n}^{I} = \beta_{0t_n} + \beta_{\xi_{int}t_n}\xi_{int}^{II} + \beta_{\xi_{lin}t_1}\xi_{lin}^{II} + \zeta_{t_n}
\end{aligned}
\tag{2}
$$

where $\beta_{t_{i-1}t_i}$ that represent the carry-over effects [12]; [6]. A sizeable positive effect means that the individuals' estimations of the construct remain stable over time [6]. In contrast, a small effect means that there has been a substantial reshuffling of the individuals' standings on the construct over time [21]. Finally, a sizeable negative effect means that there has been a reversal of the position of individuals on the structure over time. $\beta_{t_{i-1}t_i}$ contributes to explaining the variability at t time.

As in the CB-SEM framework, the model must be evaluated: first the measurement model and then the structural model. For the measurement model the Dillon-Goldstein's Rho, the mean of communalities and the mean redundancies must be examined. The structural model quality of the inner model must be assessed by examining the following indices: the regression weights, the coefficient of determination (R^2), the redundancy index, and the goodness-of-fit (GoF) statistics [24]. If the structural model quality is well assessed, but one or more *carry-over effects* are negative, this means there are two or more subsamples, with different growth curves. In this case, we suggest splitting the sample into two or more subsamples.

Subsequently, multi-group comparisons could be used to test any differences in the structural path estimates.

3 Conclusions

In this paper we have showed how the PLS-PM approach can be successfully used to analyze longitudinal data. Using PLS-PM we have the best estimation of the measurement model without any problem concerning its identification. The choice of using the PLS-PM is particularly useful for several reasons. First of all, this approach is applicable to small samples and it is capable of estimating rather complex models (with many latent and observable variables) with less restrictive requirements concerning normality and variable and error distributions [10]. Furthermore, PLS-PM approach provides the possibility of working with missing data and in the presence of multi-collinearity. Another advantage of this approach, as compared to other multivariate techniques, is that it examines simultaneous a series of dependence relationship, using a single statistical approach to test the full scope of projected relations [10].

References

1. Ancona, D.G., Okhuysen, G. A and Perlow, L. A: Taking time to integrate temporal research. *Academy of Management Review*, 26 (4), 512-529 (2001)
2. Baltes, P. B. and Nesselroade, J. R: The developmental analysis of individual differences on multiple measures. Academic press (1973)
3. Braun, M. T., Kuljanin, G., and DeShon, R. P.: Spurious results in the analysis of longitudinal data in organizational research. *Organizational Research Methods*, 16(2), 302-330. (2013)
4. Cataldo, R., Grassia, M.G., Lauro, N,C., and Marino, M.:Developments in Higher-Order PLS-PM for the building of a system of Composite Indicators. *Quality & Quantity*, 51 (2), 657–674, Springer (2017)
5. Collins, L. M., and Sayer, A. G.: New methods for the analysis of change. emphAmerican Psychological Association. (2001)
6. Duncan, T.E.,Duncan, S.C. and Strycker, L.A.:An introduction to latent variable growth curve modeling: Concepts, issues, and application, Routledge. 2013)
7. Esposito, Vinzi, V., Chin, W. W., Henseler, J., Wang, H.: *Handbook of Partial Least Squares (PLS): Concepts, Methods and Applications*, Springer, Berlin, Heidelberg, New York (2010)
8. Fornell, C. and Cha, J.:Advanced Methods of Marketing Research, ed. RP Bagozzi, Blackwell, Cambridge. (1994)
9. Henseler, J. and Chin, W. W: A comparison of approaches for the analysis of interaction effects between latent variables using partial least squares path modeling, *Structural Equation Modeling*, 17 (1), 82-109 (2010)
10. Henseler, J., Ringle, C.M. and Sinkovics, R.R.:The use of partial least squares path modeling in international marketing. *New challenges to international marketing*, 277–319, Emerald Group Publishing Limited (2009)
11. Kessler, R. C: Greenberg. DF (1981). Linear panel analysis: Models of quantitative change, New York: Academic Press (1986)

12. Johnson, M.D., Herrmann, A. and Huber, F.:The evolution of loyalty intentions, *Journal of marketing*, 70 (2), 122-132, SAGE Publications Sage CA: Los Angeles, CA. (2006)
13. Jones, E., Sundaram, S.and Chin, W.:Factors leading to sales force automation use: A longitudinal analysis. *Journal of personal selling & sales management*, 22 (3), 145-156, Taylor & Francis. (2002)
14. Joreskog, K. G. and Van Thillo, M.: LISREL: A general computer program for estimating a linear structural equation system involving multiple indicators of unmeasured variables. ERIC. (1972)
15. Lauro, N.C., Grassia, M.G. and Cataldo, R.:Model based composite indicators: New developments in partial least squares-path modeling for the building of different types of composite indicators. *Social Indicators Research*, 135 (2), 421–455, Springer (2018)
16. Lohmöller, J.B.:Latent variable path modeling with partial least squares, Springer Science & Business Media (2013)
17. McArdle, J. J. and Nesselroade, J. R: Longitudinal data analysis using structural equation models., *American Psychological Association*. (2014)
18. McArdle, J.J.: Dynamic but structural equation modeling of repeated measures data. *Handbook of multivariate experimental psychology*, 561–614. (1988)
19. Rajala, R. and Westerlund, M.:Antecedents to Consumers' Acceptance of Mobile Advertisements-A Hierarchical Construct PLS Structural Equation Model, 2010 43rd Hawaii International Conference on System Sciences, 1–10, IEEE (2010)
20. Roemer, E.:A tutorial on the use of PLS path modeling in longitudinal studies, *Industrial Management & Data Systems*, 116 (9), 1901-1921,Emerald Group Publishing Limited. (2016)
21. Selig, J.P. and Little, T.D.:Autoregressive and cross-lagged panel analysis for longitudinal data. (2012)
22. Shea, C.M. and Howell, J.M.:Efficacy-performance spirals: An empirical test, *Journal of Management*, 26 (4), 791-812, Sage Publications Sage CA: Thousand Oaks, CA (2000)
23. Stoel, R.D., van den Wittenboer, G. and Hox, J.:Methodological issues in the application of the latent growth curve model, *Recent developments on structural equation models*, 241–261, Springer (2004)
24. Tenenhaus, M., Esposito Vinzi, V., Chatelin, Y. M., Lauro, C. N.: PLS Path Modeling. *Computational Statistics and Data Analysis*, 48 (1), 159 - 205 (2005)
25. Wilson, B.:Using PLS to investigate interaction effects between higher order branding constructs. *Handbook of partial least squares*, 621–652, Springer (2010)
26. Wold, H.:Path models with latent variables: The NIPALS approach. *Quantitative sociology*, 307-357, Elsevier (1975)
27. Wold, H.:Soft modeling: the basic design and some extensions. *Systems under indirect observation*, 2, 343 (1982)
28. Wold, H.:Partial least squares. S. Kotz and NL Johnson (Eds.), Encyclopedia of statistical sciences (vol. 6), Wiley, New York (1985)