Electronic Journal of Applied Statistical Analysis
EJASA (2012), Electron. J. App. Stat. Anal., Vol. 5, Issue 3, 458 – 464
e-ISSN 2070-5948, DOI 10.1285/i20705948v5n3p458

GENERAL DISTRESS AS SECOND ORDER LATENT VARIABLE ESTIMATED THROUGH PLS-PM APPROACH

Enrico Ciavolino*

Department of History, Society and Human Studies, University of Salento, Italy

Received 29 July 2012; Accepted 13 November 2012
Available online 16 November 2012

Abstract: *The aim of the paper is to present a Partial Least Squares Path Modeling based on high-order latent variables to analyze the concept of General Distress. In order to depict General Distress concept, two approaches to second order construct modeling are presented and compared: the repeated indicators and the two-step approach. It is shown how to implement the two estimation approaches and a comparative study is proposed.*

Keywords: *High-order construct model, Partial Least Squares Path Modeling, psychological assessment.*

1. Introduction

Partial least squares (PLS) enables researchers in many field of social sciences to investigate models at high level of abstraction. The dimensions of a higher-order construct could be conceptualized under an overall abstraction, and it is theoretically meaningful to use this abstraction for the representation of the dimensions, instead of merely interrelating them.
The General Distress concept is analyzed trough the Partial Least Squares Path Modeling (PLS-PM) in order to show the method implementation and to compare the two main approaches available in literature: the *repeated indicators* and the *two-step* approach. The first one has been proposed by Lohmöller (1989) [9], who suggested to built up the higher-order LV as a general construct linked to all the manifest variables (MVs) of the lower-order constructs. In this way, each item in the model has direct relations both with the specific LVs, at the lower level, and with the general LV, at the higher level in the model.

* E-mail: enrico.ciavolino@unisalento.it

In the second one, the LV scores are initially estimated in a model without second-order constructs. The estimated LVs scores are subsequently used as indicators in a separate higher-order structural model analysis [1].

The work is articulated as follows: in *section 2* The PLS-PM extensions to higher-order construct modeling are described; *section 3* presents the empirical evidence of General Distress modeled as a second-order latent variable; in *section 4*, results from the analysis and conclusions are reported.

2. Estimating higher-order construct model through PLS-PM

The utility of higher (or hierarchical) construct models is based on theoretical and empirical considerations [4]. For many authors, this kind of models allows for a reduction of the model complexity and theoretical parsimony: theory often adopt broad concepts consisting of specific dimensions. As stated by Gorsuch [6], "factors are concerned with narrow areas of generalization where the accuracy is great [whereas] higher-order factors reduce accuracy for an increase in the breadth of generalization". Matching the two levels of abstraction leads to the detection of two sources of influence on every single measures: while the first-order accounts for the unique variance of measures in one of the specific LVs, the second-order contains the common variance as represented by the general construct.

In addition to these theoretical grounds, empirical reasons lead to prefer a hierarchical structure for the data. One issue concerns the internal consistency, namely the information about the equivalence, or homogeneity, of the MVs (usually the test items). Indeed, broader concepts will include heterogeneous items, thus showing low internal consistency. Finally, the multidimensional construct at issue will contain large amounts of specific and group variance [8] and this will affect the construct validity.

The following two approaches are available within the frame of the PLS-PM in order to built up and estimate second-order LV models.

2.1 The repeated indicators approach modeling

In 1989 Lohmöller [9], in his book about the development of SEM in the frame of the PLS-PM, proposed a special procedure for the case of hierarchical constructs, which he called Hierarchical Component Model, also known as the Repeated Indicators Approach. Such an approach is the most popular when estimating higher-order constructs through PLS. The procedure is very simple: "a second-order factor is directly measured by observed variables for all the first-order factors. While this approach repeats the number of manifest variables (MVs) used, the model can be estimated by the standard PLS algorithm" [12].

Repeated indicator approach modeling can be specified by considering the following three equations:

$$\xi^{I}_{m,1} = \mathbf{B}_{m,q} \cdot \xi^{II}_{q,1} + \zeta_{m,1} \quad (1)$$

$$\mathbf{x}_{p,1} = \Lambda^{I}_{p,m} \cdot \xi^{I}_{m,1} + \delta_{p,1} \quad (2)$$

$$\mathbf{x}_{p,1} = \Lambda^{II}_{p,1} \cdot \xi^{II}_{1,1} + \varepsilon_{p,1} \quad (3)$$

where the subscripts m and p are the number of, respectively, the first-order LVs and the manifest variables (MVs) in the model, the subscript q is the number of second-order LV. The vectors $\boldsymbol{\xi}^{I}$, $\boldsymbol{\xi}^{II}$, $\mathbf{x}$, $\boldsymbol{\zeta}$, $\boldsymbol{\delta}$ and $\boldsymbol{\varepsilon}$, indicate respectively the first and the second order LVs, the MVs, the structural and the measurement errors terms. The matrices $\mathbf{B}$, $\boldsymbol{\Lambda}^{I}$, $\boldsymbol{\Lambda}^{II}$, define the path coefficients linking the LVs and the factor loading linking, respectively, the MVs to the first and second-order LVs. The structural or inner model (equation 1) specifies the relationships among the first and the second order latent variables (LVs.). The equations 2 and 3 denote the measurement models, where the MVs, measuring each first-order LV, are repeated in order to represent the higher-order construct.

Because of its simplicity, this approach is the most used by researchers who want to model higher-order constructs with PLS [13, 10]. A possible biasing of the estimates by relating variables of the same type together through PLS estimation should be taken into account. Moreover, according with Rajala *et al.* [11], the Repeated Indicators approach could be applied, provided that all the measurement relationships (between the LVs and the related observed variables) are modeled as reflective. Formative relations from the first-order to second-order LVs can also be hypothesized (aggregate model), as shown in different studies [5, 10].

2.2 The two-step approach modeling

Another way of building a higher-order model is the Two-Step Approach: the LV scores are initially estimated in a model without the second-order LV [1]. Once the first-order LV scores are computed, they are subsequently used as indicators in a separate higher-order structural model analysis. The first-order specific LVs are then a linear combination of the high-order general LV, while the observed variables are directly related only to the specific dimensions. Such an approach may offer advantages when estimating higher-order models with formative indicators [12].

The implementation is not one simultaneous PLS run; this implies that any second-order LV, estimated in stage two, is not taken into account when estimating LV scores at the first stage. The first step of estimation is made by considering just the measurement model which provides the estimation of the first-order LVs, as reported in the following equation:

$$\mathbf{x}_{p,1} = \Lambda^{I}_{p,m} \cdot \xi^{I}_{m,1} + \delta_{p,1} \qquad (4)$$

In the second step, the estimated scores $\hat{\xi}^{I}$, obtained in the first step, are used as indicator of the second order LV:

$$\hat{\xi}^{I}_{m,1} = \mathbf{B}_{m,1} \cdot \xi^{II}_{1,1} + \zeta_{m,1} \qquad (5)$$

The main difference in the results of the two approaches described above is in the *directedness* of the impact of the second-order LV on the observed variables. While the repeated indicators approach links directly the second-order LV both to the first-order LVs and the manifest indicators, in the two-step estimation the general construct has direct effects on the sub-dimensions and only indirect effects on the MVs.

3. Measuring the General Distress as second-order LV

In our empirical illustration we use the concept of General Distress in order to show how a second-order LV structural equation model can be estimated through the PLS-PM approach. The high-order model presented is reflective both at the measurement and at the structural level.

The construct of General Distress here hypothesized arises from the analysis of the structure of the *Outcome Questionnaire 45* (OQ-45) [7]. The OQ-45 is a psychological assessment instrument designed to assess the "patient functioning" in three domains, namely:

- *Symptom Distress* (SD): 25 items about the most common disorders in mental health care (depression, anxiety and addiction to alcohol and drugs).
- *Interpersonal Relations* (IR): 11 items measuring functioning of the patient in his/her relationships with partner, family and friends.
- *Social Role* (SR) performance: 9 items on the functioning in the school/work and leisure contexts.

The questionnaire consists of a total of 45 items measured through a 5 point Likert-type scale. Scores on the whole scale as well as on each of the three sub-scales are used to track the patient change along the therapy. The questionnaire is reported in the supplementary file.

The data used for the analysis are those collected by De Jong, Nugter, Polak, Wagenborg, Spinhoven and Heiser [3] for their study aimed at assessing the construct validity of the OQ-45 in the Dutch population (n= 2129).

Based on the results of the authors analysis, we use the four-factor solution as the starting point of our analysis and performed the estimation of the second-order model with the two approaches. In these models, a general distress construct is hypothesized to underlie the first-order dimensions investigated by the OQ-45.

3.1 Estimation of the second-order LV model

The two approaches to the estimation of second-order LV models have been implemented in order to measure the General Distress construct.

In Table 1 the structural coefficients, linking the general distress to the specific sub-dimensions, are reported. The structural relationships estimated by the repeated indicators show the OQ-45 is strongly affected by a more general dimension of distress, while, the two-step approach produces lower estimates, but nevertheless it reproduces the same ordering of intensities.

The dimension which is most affected by the general distress is the *Symptom Distress scale*, while the less influenced is the *Social Role* dimension, inherently with the fact that the OQ-45 questionnaire is a tool for assessing the clinical disease.

Figure 1 shows the direct effects of the overall distress on the observed items. Since in the two-step approach no direct links are provided between the second-order LV and the MVs, the direct effect has been computed as the product of the structural path coefficients and the first-order measurement model's weights. As an example, the effect of General Distress on the third item, SD23, results from the product of the path coefficient General Distress → Symptom Distress ($\gamma_{1,1}$=0.545) and the measurement weight Symptom Distress → SD23 ($\lambda^{I}_{1,23}$=0.313), so that the effect on the item is $\lambda^{II}_{1,23}=\gamma_{1,1}\cdot\lambda^{I}_{1,23}=0.545\cdot0.313=0.170$.

The differences between the two-step and the repeated indicators are particularly great for the construct Social Role, especially for the items SR21, SR38 and SR44. The other constructs produce substantially the same effects on MVs when estimated with both methods.

Table 1. Links between the first and the second layer.

	General Distress	
LVs	Repeated Indicators	Two Step
Symptom Distress	0.960	0.545
Anxiety and Somatic Distress	0.981	0.504
Interpersonal Relations	0.856	0.502
Social Role	0.771	0.444

Most of the highest second-order effects fall in the domain of Symptom Distress. Conversely, items less affected by the general distress are those related to Interpersonal Relations and Social Role.

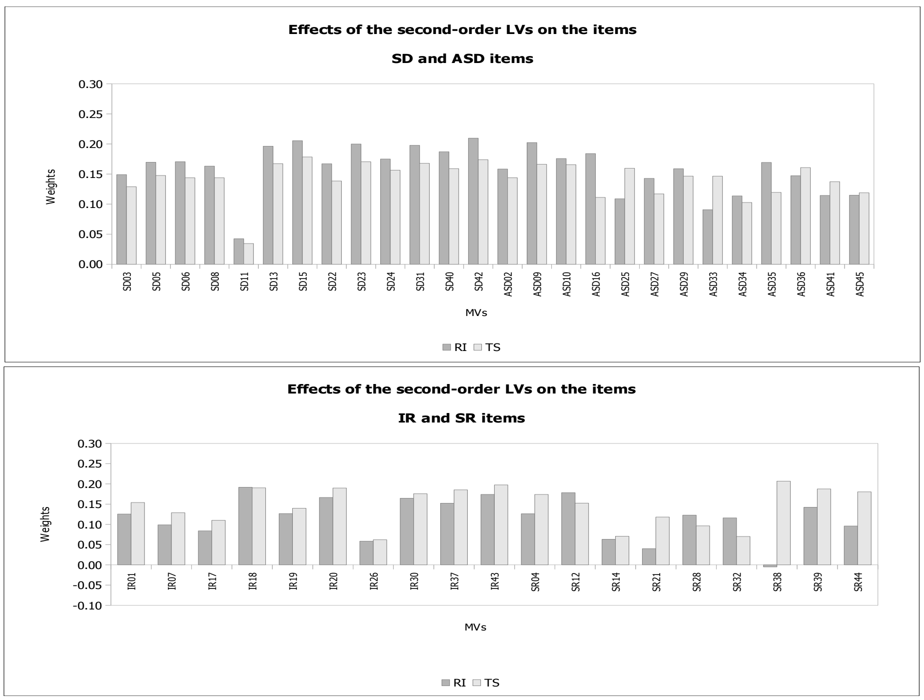

Figure 1. Effects of the second-order LV on the items.

Table 2 reports the reliability measures assessing the psychometric properties of the second-order LV and the index assessing the goodness of fit of the two models. Both Cronbach's alpha and composite reliability are more than satisfactory in both models.

Table 2. Reliability of the second-order LV and overall goodness of fit index.

	Cronbach's Alfa	***Composite Reliability***	***Communality***	***GoF***
Repeated Indicator	0.947	0.953	0.333	0.537
Two-Step	0.879	0.918	0.737	0.672

As a matter of fact, the amount of variability of the MVs captured by the general distress construct is very scarce when the repeated indicators is adopted; conversely, communality is good when the general distress is measured through the two-step. It should be noted that the low value for communality obtained with the repeated indicators is due to the fact that the second-order LV is measured by the collection of several heterogeneous items, and this negatively affect the construct's internal consistency. The right part of Table 2 reports an index proposed by Amato *et al.* [2] as a global fit measure for PLS-PM. Both approaches produce values that exceed the cut-off value of 0.36 for large effect sizes [13], so the two second-order models have a good fit. Finally, the proportion of variance explained by each layer in the second-order model is reported in Table 3. For both approaches, more than half of the variability is captured by the first layer of the model. The variance explained by the second layer is higher when the repeated indicators approach is used.

Table 3. Variance explained by the three models.

	Explained Variance	
	Layer 2	***Layer 1***
Repeated Indicator	0.439	0.561
Two-Step	0.301	0.699

4. Conclusions

The purpose of this paper was to provide researchers using PLS-PM for analysing high-order LVs with an overview of the two main approaches present in the literature, as well as an evaluation of their suitability based on an empirical evidence. The two approaches presented are the repeated indicators and the two-step approach.

In order to show the application of the two approaches and compare their results, a case study in the field of psychological assessment has been analyzed. The main difference concerns the amount of variability explained by the second-order constructs, as expressed by communality and proportion of variance explained in the model. In fact, values for the communality, which is the variability in the block of items which is explained by the LV, are the lowest when the repeated indicators approach is used (33.3%, against the 73.7% obtained by the two-step approach). Such a difference is not surprising, since the second-order construct built up with the repeated indicators approach is measured by 45 items grouped into related, but heterogeneous, dimensions. This estimation feature could lead the researcher to prefer the two-step estimation approach. Conversely, the proportion of variance explained by the general constructs with respect to the total variance explained by the model is reduced from 43.9% provided by the repeated indicators to 30.1% of the two-step approach modeling. A crucial question for the researcher wanting to model a high-order constructs is which, between the two approaches presented, leads to best results. We believe that none of the two approaches is the best and that they are alternative, rather than competitive, approaches for the estimation of second-order constructs. The choice will then depend on the meaning expressed by the high-order dimension investigated and which is its role with respect to the lower-order dimensions. The interpretation of the case study can help to understand the differences of the two approaches in terms of application perspective. With the repeated indicators approach, the General Distress defines a broad psychological dimension expressing the total amount of patients distress. The model's path coefficients define the intensity of the causal relationships between the General Distress and its sub-dimensions, represented by first-order LVs. Which means, for instance, keeping the other parameters constant, if the General Distress increases of one unit, it will affect the interpersonal relations of 0.856. In the two-step approach, the General Distress defines again a global measure of patients' distress, but in this case the relationships with the specific dimensions consist of structural coefficients of a measurement model. The path coefficients (0.277; 0.624; 0.725) reflect the composition of the General Distress; they indeed do not represent how much the high-order dimension affects the first-order sub-dimensions, but the extent to which the first-order constructs reflect the higher level of abstraction.

References

[1]. Agarwal, R., Karahanna, E. (2000). Time Flies when you are Having Fun: Cognitive Absorption and Beliefs about Information Technology Usage. *MIS Quarterly*, 24(4), 694.

[2]. Amato, S., Esposito Vinzi, V., Tenenhaus, M., 2004. A global goodness-of-t index for PLS structural equation modeling. *Oral Communication to PLS Club*, HEC School of Management, France, March 24.

[3]. De Jong, K.D., Nugter, M.A., Polak, M.G., Wagenborg, J.E.A., Spinhoven, P., & Heiser, W. J. (2007). Questionnaire (OQ-45) in a Dutch Population : A Cross-Cultural Validation. *Clinical Psychology*, 301, 288–301.

[4]. Edwards, Jeffrey R. (2001). Multidimensional Constructs in Organizational Behavior Research : An Integrative Analytical Framework. *Organizational Research Methods*, 4(2), 144–192.

[5]. Esposito Vinzi, V., Chin, W., Henseler, J.,Wang, H. (2010). *Handbook of Partial Least Squares - Concepts, Methods and Applications*. Berlin: Springer.

[6]. Gorsuch, R. L. (1983). *Factor Analysis 2nd Edition*. Hillsdale: Psychology Press.

[7]. Lambert, M.J., Burlingame, G.M., Umphress, V., Hansen, N.B., Vermeersch, D.A., Clouse, G.C., Yanchar, S.C. (1996). The Reliability and Validity of the Outcome Questionnaire. *Clinical Psychology & Psychotherapy*, 3(4), 249–258.

[8]. Law, K.S., Wong, C., Mobley, W.H. (1998). Toward a taxonomy of multidimensional constructs. *Academy of Management Review*, 23(4), 741–755.

[9]. Lohmöller, J.-B. (1989). *Latent variable path modeling with partial least squares*. Heidelberg: Physica-Verlag.

[10]. Martinez-Ruiz, A., Aluja-Banet, T. (2009). Toward the definition of a structural equation model of patent value: PLS path modelling with formative constructs. *REVSTAT Statistical Journal*, 7 (3), 265-290.

[11]. Rajala, R., Westerlund, M. (2010). Antecedents to Consumers' Acceptance of Mobile Advertisements – A Hierarchical Construct PLS Structural Equation Model. *Proceedings of the 43rd Hawaii International Conference on System Sciences* - 2010 (pp. 1–10).

[12]. Reinartz, W., Krafft, M., Hoyer, W.D. (2004). The Customer Relationship Management Process : Its Measurement and Impact on Performance. *Journal of Marketing Research*, XLI(August), 293–305.

[13]. Wetzels, M. Odekerken-Schroder, G. Van Oppen, C. (2009). Using PLS Path Modeling for Assessing Hierarchical Construct Models: Guidelines and Empirical Illustration. *MIS Quarterly*, 33 (1), 177-195.