Electronic Journal of Applied Statistical Analysis
EJASA (2012), Electron. J. App. Stat. Anal., Vol. 5, Issue 3, 333 – 338
e-ISSN 2070-5948, DOI 10.1285/i20705948v5n3p333

INCONSISTENCIES OF THE PLS-PM APPROACH TO STRUCTURAL EQUATION MODELS WITH FORMATIVE-REFLECTIVE SCHEMES

Marco Fattore[*] [(1)], Matteo M. Pelagatti[(2)], Giorgio Vittadini[(1)]

[(1)] *Department of Statistics and Quantitative Methods, University of Milan-Bicocca, Italy*
[(2)] *Department of Economics, Quantitative Methods and Business Strategies, University of Milan-Bicocca, Italy*

Received 13 July 2012; Accepted 07 October 2012
Available online 16 November 2012

Abstract: *The aim of the paper is to show evidence of possible inconsistencies of PLS-PM as a statistical tool to estimate structural equation models with formative-reflective schemes. We pursue this goal, discussing a real-data example where PLS-PM fails to identify existing causal links between the variables concerned. We also suggest a possible formal interpretation of this fact, focusing on the way PLS-PM represents the link between reflective latent variables and their indicators.*

Keywords: *Formative-reflective models, PLS-PM, Causal models.*

1. Introduction

This paper discusses empirical evidence of possible limitations of the PLS Path Modeling (PLS-PM) algorithm as a statistical tool to estimate structural equation models with formative-reflective schemes. Many algorithms and methodologies exist to address the estimation of structural models with both formative and reflective latent constructs, the most frequently applied being the Lisrel and the PLS-PM tools. The drawbacks of the Lisrel approach have been clearly understood and extensively discussed (see for instance [11] and [12]), and in many cases the PLS-PM approach is preferred. PLS-PM suffers no indeterminacy problems, it requires no assumption of multinormality in the data and it is efficiently implemented in many software packages. Although it is extensively applied to formative-reflective models and also to purely reflective models (see for example [5]), PLS-PM formally represents latent variables (LVs) in a formative way. As a consequence, we argue that applying PLS-PM to models comprising reflective latent variables may lead to logical and interpretational inconsistencies (see also [7]

[*] E-mail: marco.fattore@unimib.it

and [12]). In the following, we give explicit evidence of these issues, commenting on a real-data example pertaining to customer equity management.

2. Technical preliminaries

In this Section, we focus on structural equation models with a formative-reflective scheme, since this is the case where PLS-PM reveals its main drawbacks and since the example discussed in the next section employs it to model the data (for more complete treatments of the topics, see [6], [9] and [12]). In a formative-reflective model, a set of p blocks of formative manifest variables (MVs), $x_{(1)},\dots,x_{(p)}$, forms p exogenous latent variables (LVs), $\xi_1,\dots,\xi_p$, which in turn form a set of q endogenous LVs, $\eta_1,\dots,\eta_q$, reflected by a set of q blocks of manifest indicators, $y_{(1)},\dots,y_{(q)}$ (for a graphical representation see Figure 1).

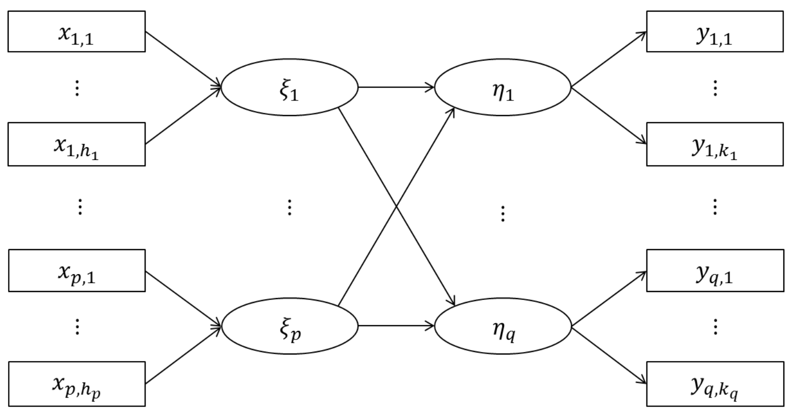

Figure 1. Path diagram of the formative-reflective model.

Omitting error in equations (which are not considered in PLS-PM), the model is specified by the following equations:

$$\boldsymbol{\xi} = \boldsymbol{\Omega}\mathbf{x},$$
$$\mathbf{y} = \boldsymbol{\Lambda}\boldsymbol{\eta} + \boldsymbol{\varepsilon},$$
$$\boldsymbol{\eta} = \boldsymbol{\Gamma}\boldsymbol{\xi},$$

where $\mathbf{x}$, $\mathbf{y}$, $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are vectors obtained by stacking the corresponding manifest and latent variables, $\boldsymbol{\varepsilon}$ is a vector of residuals and $\boldsymbol{\Omega}$, $\boldsymbol{\Lambda}$ and $\boldsymbol{\Gamma}$ are suitable matrices. PLS-PM is a limited information method which applies an iterative algorithm that separately estimates the several blocks of the measurement model and then, in a second step, estimates the structural model coefficients. The estimation of the latent variable scores is obtained through the alternation of the outer and the inner estimations, iterating until convergence. Although there is no room here to describe the algorithm in detail (see [9] for an up-to-date treatment), it is worth mentioning that it produces estimated scores for ξ_i ($i = 1,\dots,p$) and η_j ($j = 1,\dots,q$), which are linear combinations of their own manifest variables. Even though in [10] (p. 52) it is stated that this fact "[…] does not affect the direction of the relationship between the latent variable and its own manifest variables

[…]", with reflective LVs this is inconsistent with the causal structure of the model (see [2]) and it is likely to be one of the roots of the drawbacks of PLS-PM illustrated in Section 3.

3. Inconsistencies of PLS-PM: empirical evidence

To illustrate possible drawbacks of PLS-PM when reflective constructs are part of the model, we comment on a real-data application provided by Bruhn, Georgi and Hadwick [3] who made the correlation matrix of their variables publicly available (Table 3 of [3]). The goal of Bruhn et al. is evaluating how intensively firms orient their customer management towards customer value and equity. They focus on the concept of *customer equity management* (CEM) defined as "all [those business] activities that aim explicitly to maximize customer equity." The model of Bruhn et al. adopts a formative-reflective scheme (see Figure 2) with three exogenous LVs (CE Analysis, CE Strategy, CE Actions) and one endogenous LV (CE Management). In the terminology of Jarvis [6], this is a type IV model (interestingly enough, quite an uncommon model in real applications).

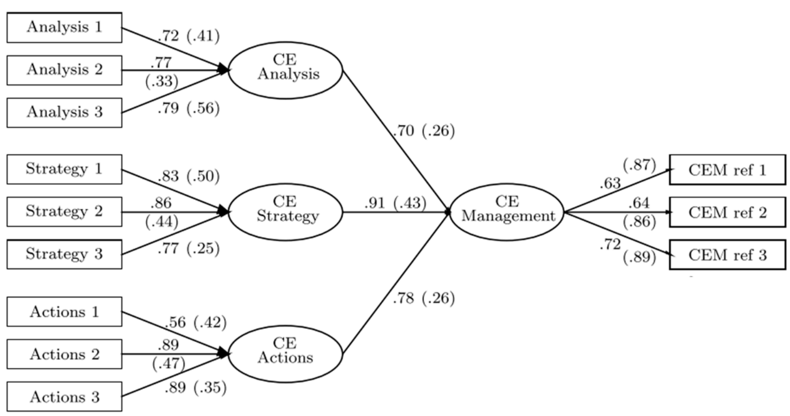

Figure 2. Model for customer equity management, reported from [3]. Output of PLS (in parenthesis) and of least squares (out of parenthesis).

The main interest of the study of Bruhn et al. is to identify causal links between the two sides of the model, in order to determine how to improve the effectiveness of customer equity management. PLS-PM is applied to estimate the model. The results are displayed in Figure 2, in parentheses. A direct check shows that (1) the correlation between the formative MVs and the exogenous LVs are quite small, ranging from 0.25 to 0.56; (2) the correlations between the exogenous LVs and the endogenous LV are similarly small, ranging from 0.26 to 0.43 and (3) the correlations between the endogenous LV and its manifest indicators are instead much stronger, ranging from 0.86 to 0.89. As a result, no neat causal path seems to exist, linking the formative side and the reflective side of the model. The endogenous LV is indeed a good predictor of its manifest indicators, but its links to the left side of the model are not clear. A manager looking at these results would conclude that CE Analysis, Strategy and Actions do not affect CE Management much (somehow a strange conclusion, we can say). To check whether the results generated using PLS-PM are reliable, we have analyzed the formative-reflective model using a simpler constrained least squares approach (see [4], for more details). More precisely:

1. Each exogenous LV has been defined as a linear combination of its own formative MVs.
2. The endogenous LV has been defined as a linear combination of the exogenous LVs.
3. Running a quasi-Newton algorithm with numerical derivatives, the coefficients of the linear combinations have been determined so as to maximize the explained variance of the reflective MVs.
4. Finally, the correlations between the manifest and the latent variables and between the exogenous and the endogenous LVs have been computed.

Formally, the above procedure minimizes the following objective function:

$$L(\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_3,\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2,\boldsymbol{\lambda}_3,\boldsymbol{\Omega})=\sum_{j=1}^{3}\frac{Tr\{E\lfloor(\mathbf{y}_j-\boldsymbol{\lambda}_j\boldsymbol{\gamma}_j'\boldsymbol{\Omega}\mathbf{x})(\mathbf{y}_j-\boldsymbol{\lambda}_j\boldsymbol{\gamma}_j'\boldsymbol{\Omega}\mathbf{x})'\rfloor\}}{Tr\{E[\mathbf{y}_j\mathbf{y}_j']\}}$$

where $\boldsymbol{\gamma}_j'$ is the *j*-th row of matrix $\boldsymbol{\Gamma}$ and the vectors $\boldsymbol{\lambda}_j$ are regression coefficients of $\mathbf{y}_j$ on $\boldsymbol{\gamma}_j'\boldsymbol{\Omega}\mathbf{x}$. Final results are reported in Figure 2, out of parentheses. Compared to the PLS-PM algorithm, the least squares approach produces very different outputs, in fact (1) the correlations between the formative MVs and the exogenous LVs range from 0.56 to 0.89 and are much higher than those in the PLS-PM output; (2) the correlations between the exogenous LVs and the endogenous LV range from 0.70 to 0.91, also very much higher than in the PLS-PM case and (3) the correlations between the endogenous LV and its manifest indicators range from 0.63 to 0.62. They are smaller than those from PLS-PM, but still quite high. Differently from PLS-PM, and consistently with theoretical considerations (see [1], [2] and [8]), the least squares approach reveals that a causal pattern linking the variables within the model does exist: MVs are highly correlated with the exogenous LVs, which, in turn, are highly correlated with the endogenous LV which, finally, is well reflected by the selected manifest indicators. The strong correlations among the variables make the LVs clearly interpretable in terms of antecedents and consequents and the formative and reflective sides of the model appear consistently linked. This is not the case in the PLS-PM output, where the exogenous LVs and the endogenous LV are too weakly correlated to link causally the formative and the reflective sides of the model and where the LVs are hardly interpretable. In summary, in this case study PLS-PM fails the primary goal of any estimation tool when applied to formative-reflective schemes: the identification of existing causal (or, at least, predictive) links among the variables in the model. In [10] (p. 52) it is stated that “Taking into account the regression framework of PLS Path Modeling […] the emphasis is more on the accuracy of predictions than on the accuracy of estimation.” In a sense, this enforces the evidence that PLS-PM may be not effective in some situations. From a mathematical point of view, the reason for this drawback is likely to be linked to the way PLS-PM defines the endogenous LV. In PLS-PM, the endogenous LV is expressed as a linear combination of the reflective MVs, that is, it is bound to belong to the linear subspace generated by the reflective indicators. This constraint is not consistent with the causal structure of the model. It leads to extracting an endogenous LV which is highly correlated with its indicators; at the same time, however, it prevents the algorithm to search for other possible solutions more correlated with the formative MVs and, thus, with the extracted exogenous LVs. Our conjecture (supported also by some theoretical arguments and simulation results reported in [12]) is that this feature of PLS-PM may in general affect its performances when reflective latent variables are concerned.

Depending upon the correlation structure among manifest variables, the "tension" existing between the exogenous and the endogenous LVs (which belong to different linear subspaces) may in fact result into inefficient and inconsistent outputs like those presented above.

4. Conclusions

In this paper we have provided empirical evidence of the ineffectiveness of the widely-adopted PLS-PM algorithm, when formative-reflective schemes are concerned. Realizing the limitations and the drawbacks of PLS-PM has more than a theoretical relevance and could be of great interest for applied statisticians, when we consider the customer equity example and the wrong conclusions a manager would be led to. Here, due to space limitations, we have confined ourselves to an empirical study. Indeed, more research is needed to further the analysis and definitively understand the mathematical roots of the possible failure of PLS-PM. We hope this paper may attract some research on the topic, highlighting that also mainstream statistical tools should be used with care.

References

[1]. Bell, D., Deighton, J., Reinartz, W.J., Rust, R.T., Swartz, G. (2005). Seven barriers to customer equity management. *J Serv Res,* 5(1), 77–85.

[2]. Berger, P.D., Bolton, R.N., Bowman, D., Briggs, E., Kumar, V., Parasuraman, A., Terry, C. (2002). Marketing actions and the value of customer assets: a framework for customer asset management. *J Serv Res*, 5(1), 39–45.

[3]. Bruhn, M., Georgi, D., Hadwick, K. (2008). Customer equity management as formative second order construct. *Journal of Business Research*, 61, 1292-1301.

[4]. Fattore, M., Pelagatti, M., Vittadini, G. (2012). A least squares approach to latent variables extraction in formative-reflective models. *EconPapers*, working paper from University Milano-Bicocca, 20120302.

[5]. Fornell, C., Johnson, M.D., Anderson, E.W., Cha, J., Everitt, B. (1996). The American Customer Satisfaction Index: Nature, Purpose, and Findings. *Journal of Marketing*, 60, 7-18.

[6]. Jarvis, B.K., MacKenzie, S.B., Podsakoff, P.M. (2003). A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research. *Journal of Consumer Research*, 30(2), 199-218.

[7]. Marcoulides, G.A., Chin, W., Saunders, C. (2009). A Critical Look at Partial Least Squares Modeling. *MIS Quarterly*, 33 (1), 171-175.

[8]. Payne, A., Holt, S., Frow, P. (2001). Relationship value management: exploring the integration of employee, customer and shareholder value and enterprise performance models. *J Market Manag,* 17(7/8), 785–817.

[9]. Tenenhaus, M., Esposito Vinzi, V., Chatelin, Y.M., Lauro, C. (2005). PLS path modeling. *Computational Statistics & Data Analysis*, 48, 159–205.

[10]. Vinzi, V.E., Trinchera, L., Amato S., (2010). PLS Path Modeling: From Foundations

to Recent Developments and Open Issues for Model Assessment and Improvement, in *Handbook of Partial Least Squares*, eds. E. Vinzi et al., Heidelberg: Springer-Verlag, 47-82.

[11]. Vittadini, G. (1989). Indeterminacy problems in the Lisrel model. *Multivariate Behavioral Research*, 24(4), 397-414.

[12]. Vittadini, G., Minotti, S.C., Fattore, M., Lovaglio, P.G. (2007). On the relationships among latent variables and residuals in PLS path modeling: the formative-reflective scheme. *Computational Statistics & Data Analysis*, 51, 5828-5846.