

# On the convergence of the partial least squares path modeling algorithm

Jörg Henseler

Received: 8 March 2008 / Accepted: 7 July 2009 / Published online: 1 August 2009  
© The Author(s) 2009. This article is published with open access at Springerlink.com

**Abstract** This paper adds to an important aspect of Partial Least Squares (PLS) path modeling, namely the convergence of the iterative PLS path modeling algorithm. Whilst conventional wisdom says that PLS always converges in practice, there is no formal proof for path models with more than two blocks of manifest variables. This paper presents six cases of non-convergence of the PLS path modeling algorithm. These cases were estimated using Mode A combined with the factorial scheme or the path weighting scheme, which are two popular options of the algorithm. As a conclusion, efforts to come to a proof of convergence under these schemes can be abandoned, and users of PLS should triangulate their estimation results.

**Keywords** Partial least squares path modeling · PLS · Convergence

**Mathematics Subject Classification (2000)** 62J99 Linear inference, regression · 91B82 Statistical methods; economic indices and measures · 91E45 Measurement and performance

## 1 Introduction

For decades, researchers have applied partial least squares path modeling to analyze complex relationships between latent variables. Many fields of research have

---

**Electronic supplementary material** The online version of this article (doi:[10.1007/s00180-009-0164-x](https://doi.org/10.1007/s00180-009-0164-x)) contains supplementary material, which is available to authorized users.

---

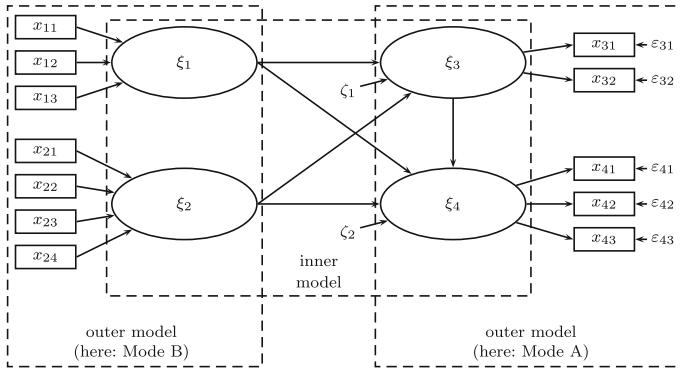
J. Henseler (✉)  
Institute for Management Research, Radboud University Nijmegen, PO Box 9108,  
6500 HK Nijmegen, The Netherlands  
e-mail: [j.henseler@fm.ru.nl](mailto:j.henseler@fm.ru.nl)

embraced the specific advantages of Partial Least Squares (PLS) path modeling, for instance behavioral sciences (cf. Bass et al. 2003) as well as many disciplines of business research, such as marketing (cf. Fornell 1992; Ulaga and Eggert 2006), strategy (cf. Hulland 1999), and management information systems (cf. Gefen and Straub 1997; Chin et al. 2003). Moreover, PLS path modeling has become a valuable tool in corporate practice, as shown by the use of PLS path modeling in companies such as L'Oréal and GfK.

The popularity of PLS among scientists and practitioners results from four genuine advantages: Firstly, PLS path modeling can be used when distributions are highly skewed (Bagozzi and Yi 1994), because “there are no distributional requirements” (Fornell and Bookstein 1982). Herman O. A. Wold, who developed PLS path modeling, positioned it as “soft modeling” as opposed to covariance-based structural equation modeling, which relies on rather “hard” assumptions. Secondly, PLS path modeling can be used to estimate relationships between latent variables with several indicators when sample size is small (Chin and Newsted 1999). As the PLS path modeling algorithm consists of OLS regressions for separate subparts of the focal path model, the complexity of the overall model hardly influences sample size requirements. Thirdly, modern easy-to-use PLS software with graphical user-interface, like SmartPLS (Ringle et al. 2007) and PLS-Graph (Soft Modeling, Inc 1992–2002), have contributed to the attractiveness of PLS. Fourthly, PLS is preferred over covariance-based structural equation modeling when improper or non-convergent results are likely (so called heywood cases, cf. Krijnen et al. 1998), as for instance in more complex models, when the number of latent and manifest variables is high in relation to the number of observations, and the number of indicators per latent variable is low. The idea that PLS “ensures against improper solutions” (Fornell and Bookstein 1982, p. 440) goes along with a strong belief among researchers: the belief that the PLS path modeling algorithm converges. For instance, Hanafi (2007, p. 280) states that “convergence [...] is always verified in practice”. Tenenhaus et al. (2005, p. 169) sum up the current knowledge about the convergence of the PLS path modeling algorithm by saying that convergence is “guaranteed only for the two-block case, but practically always encountered in practice even with more than two blocks”. Even though the PLS path modeling algorithm may converge in practice, there is concern about the missing proof of its convergence (cf. Hwang and Takane 2004), inspiring researchers to search for this proof.

In this contribution, I show that it will be impossible to deliver a general proof for the convergence of the PLS path modeling algorithm for more complex models. More specifically, I will present several datasets for which the PLS path modeling algorithm does not converge.

This paper is structured as follows. Section 2 provides a brief introduction into the PLS path modeling algorithm. Section 3 summarizes the literature on the convergence of the PLS path modeling algorithm. Section 4 describes the data generation procedure by means of which six datasets were found that led to a non-convergence of the PLS path modeling algorithm. Moreover, it reports how the path model was estimated precisely, and how the algorithm behaved in the six cases. Finally, Sect. 5 draws conclusions for scientists interested in the development and application of PLS path modeling as well as for users of PLS path modeling in general.



**Fig. 1** A simple PLS path model

## 2 PLS path modeling

Partial Least Squares is a family of alternating least squares algorithms, or ‘prescriptions’, which extend principal component and canonical correlation analysis. The method was designed by Herman Wold (1966, 1973, 1974, 1982, 1985a,b, 1989) for the analysis of high dimensional data in a low-structure environment and has undergone various extensions and modifications. This paper builds on Wold (1982) basic algorithmic design and some extensions developed by Lohmöller (1989).

### 2.1 The nature of PLS path models

Partial Least Squares path models are formally defined by two sets of linear equations: the inner model and the outer model. The inner model specifies the relationships between unobserved or latent variables, while the outer model specifies the relationships between a latent variable and its observed indicators or manifest variables. However, the various streams of literature do not always employ the same terminology. For instance, publications addressing CBSEM (e.g., Rigdon 1998) often refer to “structural models” and “measurement models”; whereas those focusing on PLS path modeling (e.g., Lohmöller 1989) use the terms “inner model” and “outer model” for similar elements of the cause-effect relationship model. As this paper deals with PLS, related terminology is used. Figure 1 depicts an example of a PLS path model.

Without a loss of generality, it can be assumed that latent and manifest variables are standardized so that the location parameters can be discarded in the following equations. The inner model for relationships between latent variables can be written as:

$$\mathcal{E} = \mathcal{E}B + Z, \tag{1}$$

where  $\mathcal{E}$  is the vector of latent variables,  $B$  denotes the matrix of path coefficients, and  $Z$  represents the inner model residuals. The basic PLS design assumes a recursive inner

model that is subject to predictor specification. Thus, the inner model constitutes a causal chain system (i.e. with uncorrelated residuals and without correlations between the residual term of a particular endogenous latent variable and its predictor variables). Predictor specification reduces Eq. 1 to:

$$E(\mathcal{E}|\mathcal{E}) = \mathcal{E}B. \quad (2)$$

PLS path modeling includes two different kinds of measurement models: Mode A and Mode B. The choice of a certain mode is subject to theoretical reasoning and in most cases results from a decision to define a measurement model as reflective or formative (see e.g. Diamantopoulos and Winklhofer 2001).

Reflective measurement models assume causal relationships from the latent variable to the related manifest variables. Within a certain block of manifest variables, each manifest variable  $x$  is modeled as a linear function of its latent variable  $\xi$  plus a residual  $\varepsilon$ :

$$x = \lambda\xi + \varepsilon, \quad (3)$$

where  $\lambda$  represents the loading coefficient. The outer relationships are also subject to predictor specification, implying that there are no correlations between the outer residuals and the latent variable of the same block. This reduces Eq. 3 to:

$$E(x|\xi) = \lambda\xi. \quad (4)$$

Formative measurement models assume causal relationships from the manifest variables to the latent variable. For one block of manifest variables, the linear relationships are given as follows:

$$\xi = X\pi + \nu. \quad (5)$$

Predictor specification reduces Eq. 5 to:

$$E(\xi|X) = X\pi. \quad (6)$$

## 2.2 The PLS path modeling algorithm

The PLS algorithm is essentially a sequence of regressions in terms of weight vectors. The weight vectors obtained at convergence satisfy *fixed point equations*; see Dijkstra (1981, 2009) for a general analysis of such equations and ensuing convergence issues. The basic PLS algorithm, as suggested by Lohmöller (1989), includes the following three stages: (1) iterative approximation of latent variable scores, (2) estimation of outer weights, outer loadings and path coefficients, and (3) estimation of location parameters.

The debate of convergence of the PLS algorithm focuses solely on the first stage, which is the core of the PLS algorithm. According to Tenenhaus et al. (2005), the first stage of the PLS path modeling algorithm consists of four steps:

- Step #1: Outer approximation of the latent variable scores.* Outer proxies of the latent variables,  $\hat{\xi}_n^o$ , are calculated as linear combinations of their respective indicators. These outer proxies are standardized; for example, they have a mean of zero and a standard deviation of one. The weights of the linear combinations result from Step #4 of the previous iteration. When the algorithm is initialized, and no weights are yet available, any arbitrary non-trivial linear combination of indicators can serve as an outer proxy of a latent variable.
- Step #2: Estimation of the inner weights.* Inner weights are calculated for each latent variable in order to reflect how strongly the other latent variables are connected to it. There are three schemes available for determining the inner weights. Wold (cf. 1982) originally proposed the centroid scheme. Later, Lohmöller (1989) developed the factor weighting and path weighting schemes. The centroid scheme uses the sign of the correlations between a latent variable—or, more precisely, the outer proxy—and its adjacent latent variables; the factor weighting scheme uses the correlations. The path weighting scheme pays tribute to the arrow orientations in the path model. The weights of those latent variables that explain the focal latent variable are set to the regression coefficients yielded from a regression of the focal latent variable (regressant) on its latent repressor variables. The weights of those latent variables, which are explained by the focal latent variable, are determined in a similar manner as in the factor weighting scheme. Regardless of the weighting scheme, a weight of zero is assigned to all non-adjacent latent variables.
- Step #3: Inner approximation of the latent variable scores.* Inner proxies of the latent variables,  $\hat{\xi}_n^i$ , are calculated as linear combinations of the outer proxies of their respective adjacent latent variables, using the afore-determined inner weights.
- Step #4: Estimation of the outer weights.* The outer weights are calculated either as the covariances between the inner proxy of each latent variable and its indicators (in Mode A, reflective), or as the regression weights resulting from the ordinary least squares' regression of the inner proxy of each latent variable on its indicators (in Mode B, formative).

These four steps are repeated until the change in outer weights between two iterations drops below a predefined limit. The algorithm terminates after Step #1, delivering latent variable scores for all latent variables. Loadings and inner regression coefficients are then calculated in a straightforward way, given the constructed indices and using Eqs. 4 and 5. In order to determine the path coefficients, a (multiple) linear regression is conducted in respect of each endogenous latent variable. The endogenous variable's scores are regressed on the latent predictor variable scores.

### 3 Current knowledge about the convergence of the PLS path modeling algorithm

Considering that Wold developed the PLS path modeling algorithm decades ago, it is surprising that the literature on its convergence is rather scarce. Tenenhaus and Vinzi

(2005) sum up the current knowledge stating that convergence is “always verified in practice but mathematically proven only for the two-block case”. For the two-block case, PLS regression (cf. Garthwaite 1994) can play the role of PLS path modeling. In this case, not only is the convergence proven, but also the outcomes of the PLS algorithm are known (Tenenhaus et al. 2005): depending on the options chosen, the PLS path modeling algorithm delivers a generalized canonical analysis (Horst 1961; Carroll 1968), a multiple factor analysis (Escofier and Pagès 1994), or a redundancy analysis (Israels 1984; Fornell et al. 1988).

Hanafi (2007) points out that there are two main procedures of the PLS path modeling algorithm: the original procedure as invented by Wold (1982, 1985a), and a modified procedure developed by Lohmöller (1989). Lohmöller’s procedure computes the latent variable scores of each latent variable  $\xi_k^{(i+1)}$  ( $1 \leq k \leq K$ ) at iteration ( $i + 1$ ) as a function of all the latent variable scores  $\xi_k^{(i)}$  obtained during the previous iteration ( $i$ ). In contrast, Wold’s procedure relies always on the most recent information available. For calculating latent variable scores  $\xi_k^{(i+1)}$  at iteration ( $i + 1$ ) for example, the procedure uses the latent variable scores  $\xi_k^{(i+1)}$  that have been generated within the present iteration if these are available; otherwise, the latent variable scores  $\xi_k^{(i)}$  from the previous iteration are used instead. Table 1 illustrates the difference between the two procedures.

The advantage of Lohmöller’s procedure is that it can efficiently be calculated by means of matrix algebra. This is the major reason why most PLS software applications, such as SmartPLS (Ringle et al. 2007), PLS-Graph (Soft Modeling, Inc 1992–2002), SPAD (Test and Go 2006), and LVPLS [Lohmöller 1987, including its extensions PLS-GUI (Li 2005) and VisualPLS (Fu 2006)], have implemented this procedure. The advantage of Wold’s procedure is its monotony properties. As Hanafi (2007) demonstrates for the case of Mode B, Wold’s procedure is monotonically convergent. Moreover, Wold’s procedure reaches a stable solution faster. These properties of Wold’s procedure result in a better performance in terms of convergence speed compared to Lohmöller’s procedure.

In a first attempt to approach a proof of convergence of the PLS path modeling algorithm, Hanafi and Qannari (2005) present a procedure that follows the same

**Table 1** Comparison of Wold’s and Lohmöller’s procedure of the PLS path modeling algorithm for an example with three latent variables

Iteration	Lohmöller’s procedure	Wold’s procedure
$i$	...	...
	$\xi_1^{(i)}, \xi_2^{(i)}, \xi_3^{(i)}$	$\xi_1^{(i)}, \xi_2^{(i)}, \xi_3^{(i)}$
$i + 1$	$\xi_1^{(i+1)} = f(\xi_1^{(i)}, \xi_2^{(i)}, \xi_3^{(i)})$	$\xi_1^{(i+1)} = f(\xi_1^{(i)}, \xi_2^{(i)}, \xi_3^{(i)})$
	$\xi_2^{(i+1)} = f(\xi_1^{(i)}, \xi_2^{(i)}, \xi_3^{(i)})$	$\xi_2^{(i+1)} = f(\xi_1^{(i+1)}, \xi_2^{(i)}, \xi_3^{(i)})$
	$\xi_3^{(i+1)} = f(\xi_1^{(i)}, \xi_2^{(i)}, \xi_3^{(i)})$	$\xi_3^{(i+1)} = f(\xi_1^{(i+1)}, \xi_2^{(i+1)}, \xi_3^{(i)})$
	$\xi_1^{(i+1)}, \xi_2^{(i+1)}, \xi_3^{(i+1)}$	$\xi_1^{(i+1)}, \xi_2^{(i+1)}, \xi_3^{(i+1)}$

purpose as PLS path modeling with Mode B, and converges. In a second attempt, Hanafi (2007) succeeds in finding a proof of convergence of the PLS path modeling algorithm for more than two blocks of manifest variables. However, this proof is limited to the following conditions:

1. Wold's procedure is used.
2. Mode B applies to all measurement models.
3. Either the centroid scheme or the factorial scheme serves as inner weighting scheme.

For all other conditions, i.e. Lohmöller's procedure, Mode A or mixed measurement models, or the path weighting scheme, there is no proof of convergence of the PLS path modeling algorithm for more than two blocks of manifest variables.

#### 4 Six cases of non-convergence of the PLS path modeling algorithm

As part of an extensive Monte Carlo simulation study about some particular behavior of the PLS path modeling algorithm, I found six datasets that made the PLS path modeling algorithm not converge. In order to facilitate a better understanding of the nature of this data, I will describe both the underlying population model and the data generation process.

All of the six datasets stem from a population model with two exogenous latent variables  $\xi_1$  and  $\xi_2$  and one endogenous latent variable  $\xi_3$ . The direct effects have coefficients of 0.3 and 0.5, respectively. Moreover, the model contains a multiplicative interaction of the two exogenous latent variables having a coefficient of 0.3. As is common for PLS path modeling, the latent variable scores are standardized, i.e. they have a mean of zero and a variance of one. In mathematical notation, the model reads as follows:

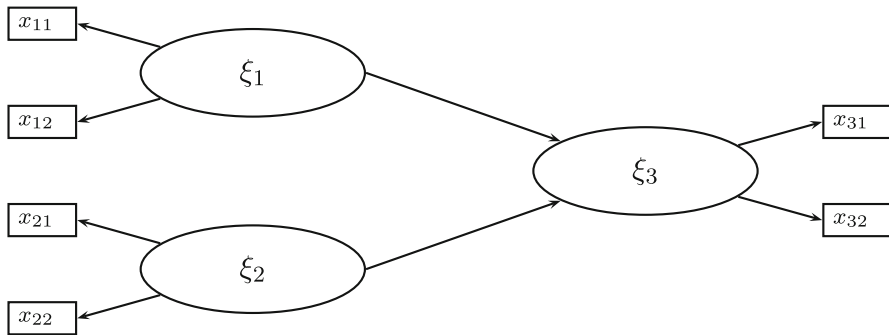
$$\xi_3 = 0.3 \cdot \xi_1 + 0.5 \cdot \xi_2 + 0.3 \cdot \xi_1 \cdot \xi_2 + \zeta \quad \xi_1, \xi_2 \sim \mathcal{N}(0, 1) \quad \zeta \sim \mathcal{N}(0, 0.57) \quad (7)$$

The latent variable scores serve as point of departure for creating reflective manifest variables. For each latent variable  $\xi_j$ , two manifest variables  $x_{j1}$  and  $x_{j2}$  are generated. All manifest variables  $x_{jh}$  have equal loadings  $\lambda = \sqrt{0.5} \approx 0.7$  and a normal-distributed unique error term:

$$x_{jh} = \lambda \cdot \xi_j + \varepsilon_{jh} \quad \varepsilon_{jh} \sim \mathcal{N}(0, 0.5) \quad \forall j \in \{1, 2, 3\}, h \in \{1, 2\} \quad (8)$$

Population parameters like those chosen are realistic for PLS path models and can be found for instance in technology acceptance studies (cf. Chin et al. 2003). For six different sample sizes, 20, 50, 100, 150, 200, and 500 observations, respectively, 500 datasets are generated. In total, this results in 3,000 different datasets characterized by the specified population model.

I analyzed the 3,000 datasets by means of PLS path modeling. However, the PLS path model for the analyses differed from the population model in that it neglected the interaction effect, i.e. only the direct effects of the exogenous latent variables  $\xi_1$  and



**Fig. 2** Estimated model

$\xi_2$  on the endogenous latent variable  $\xi_3$  were estimated. Neglecting to analyze nonlinear and in particular interaction effects is a typical and widely observed phenomenon in social and business science. Figure 2 depicts the estimated model. In the light of the estimated model, it becomes obvious that all datasets conform to the so-called “rule of ten”, which says that the number of observations should at least be ten times the number of independent variables of the most complex regression occurring in the model (Chin 1998). During the PLS estimation phase, regressions occur in the measurement model as well as in the structural model. In the present case, the structural model requires the most complex regression containing two independent variables. The recommended minimal number of observations is therefore twenty.

All computations, i.e. the simulation data generation as well as the PLS path model estimation, were done with R (version 2.5.1; R Development Core Team 2007). The used PLS path modeling algorithm is an immediate implementation of the exhaustive description provided by Tenenhaus et al. (2005). They refer to the PLS path modeling algorithm as developed by Lohmöller (1989). For all floating comma operations, I used double.

Three selected options are worth mentioning. Firstly, for the inner estimation, I applied the path weighting scheme as suggested by Lohmöller (1989). Currently, this inner weighting scheme is widely recommended because it is the only scheme that takes into account the causal order within the structural model. Secondly, as initializations of the outer estimates of the latent variables, I used the first indicator of each block as suggested by Wold (1982, p. 13) and concretized by Tenenhaus et al. (2005):

$$\xi_j = x_{j1} \quad (9)$$

Remarkably, Tenenhaus et al. (2005) point out that “[t]he starting step of the PLS algorithm consists in beginning with an arbitrary [non-trivial, the author] vector of weights,  $w_{jh}$ ”. Thirdly, the sum of absolute changes in weights from one iteration to another was tracked and compared with a threshold of  $10^{-5}$ . If it falls below this threshold, the algorithm is terminated. This particular stop criterion is suggested by Wold (1982), and often implemented in PLS path modeling software. In any case, an iteration counter assures that the algorithm would stop after a maximum of 500 iterations.



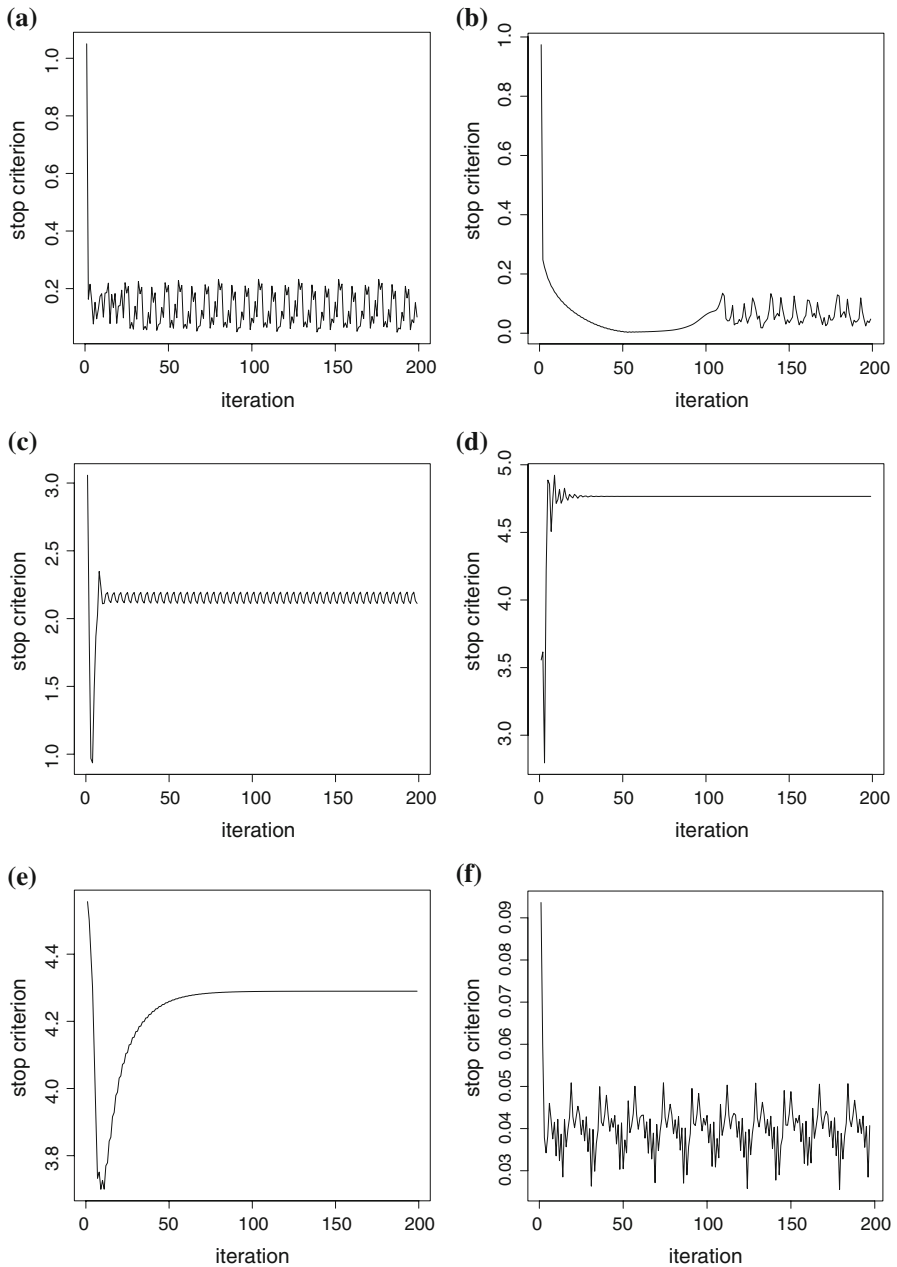
This is what happens with six datasets. That means that in six occasions the PLS path modeling algorithm does not converge within 500 iterations. Doubling the iteration limit to 1,000 does not change the result; the PLS path modeling algorithm still does not converge. This surprising finding calls for a triangulation. Two different PLS path modeling software applications, SmartPLS (version 2.0 M3, [Ringle et al. 2007](#), with a special extension for latent variable initialization) and XLSTAT-PLSPM ([Addinsoft SARL 2007–2008](#)) come to identical results, namely that for the six datasets the PLS path modeling algorithm does not converge. A closer look at the six datasets reveals that five datasets contain twenty observations, and one contains fifty observations. All of the datasets can be downloaded from the CS online repository.

In order to learn more about the generalizability of the phenomenon of non-convergence, I assessed the sensitivity of the PLS estimation in regard to changes in the weighting scheme, the latent variable initialization, and the structural model specification. First I tested whether the results would remain identical if I applied inner weighting schemes other than the path weighting scheme. Using the centroid scheme, i.e. Wold's initial suggestion, as inner weighting scheme, the PLS path modeling algorithm converges for all of the six datasets. However, applying the factorial scheme, it turned out that some of the six datasets caused non-convergence of the PLS path modeling algorithm, too. Moreover, I examined whether different latent variable initializations evoked the same outcomes. As alternative initializations of the outer estimates of the latent variables, I used the sum of the indicators of each block:

$$\xi_j = \sum_h x_{jh} \quad (10)$$

Not only is the sum of the indicators a popular proxy for latent variable scores, it is also the standard option of SmartPLS. Using the sum of the indicators, the PLS path modeling algorithm converged for all of the six datasets. Finally, I altered the specification of the structural model. I included an additional arrow pointing from  $\xi_1$  to  $\xi_2$ , in that way making the path model a complete graph. Under this condition, in all six cases convergence was achieved.

Figure 3 depicts the development of the stop criterion, i.e. the sum of absolute changes in weights (the first couple of iterations are omitted for better visibility). Obviously, the sum of absolute changes in weights does not converge towards zero for any of the six datasets. The datasets show quite a different behavior. For some of them, the stop criterion seems to converge towards a non-zero constant value. For others, the stop criterion seems to follow a periodical pattern. The second dataset probably exhibits the most worrying behavior: At the beginning, the stop criterion heads towards a local minimum (0.002889), before it rises again and enters into a phase of oscillation between values of 0.012531 and 0.135129. Besides the non-convergence, another issue should attract attention: the local minimum at the beginning. Had the stop criterion been higher, say 0.005 instead of  $10^{-5}$ , a researcher would never have known that PLS had not come to a stable solution. A hypothetical researcher would also have come to this conclusion if he or she had used a different convergence criterion, namely the sum of squared changes in weights from one iteration to another as recommended for instance by [Wold \(1982\)](#).



**Fig. 3** Development of the stop criterion for the six datasets. **a** Dataset 1: iterations 2–200; **b** Dataset 2: iterations 2–200; **c** Dataset 3: iterations 2–200; **d** Dataset 4: iterations 2–200; **e** Dataset 5: iterations 2–200; **f** Dataset 6: iterations 4–200

## 5 Discussion

In order to get a complete insight into the results of the present study, it is indispensable to be aware of its limitations. Firstly, the present study is limited in that it only provides counterexamples that use Mode A for all measurement models. Nothing can be said about the convergence of PLS path modeling when measurement models with Mode B are used. Secondly, the inner weighting scheme seems to be of importance when it comes to the issue of convergence. Whilst some models did not converge when the path weighting scheme was used, and some when the factorial scheme was used, all models did converge under the centroid scheme. Therefore, the choice of the inner weighting scheme deserves more attention from PLS researchers and users than it has received up to now in literature. Thirdly, from the perspective of graph theory, the estimated model is not a complete graph, because it does not have an arrow connection between the latent variables  $\xi_1$  and  $\xi_2$ . If the graph is completed, i.e. the missing arrow is added, the PLS path modeling algorithm (with path weighting scheme) converges. It remains open whether the PLS path modeling algorithm always converges or not if the focal PLS path model forms a complete graph. Fourthly, one cannot completely rule out the possibility that a lack of numeric precision has caused the non-convergence. Although the phenomenon of non-convergence occurred on three independent software implementations, one must recognize that all three implementations use the same numeric precision, namely double. Still, taking into account that the computational precision encompassed thirteen digits while non-convergence unfolded itself already at the third digit, the chance is very low that non-convergence happens only due to numeric imprecision. Fifthly, this study could not give an answer to the question why the algorithm has not converged. This question remains open, and requires further research. As a point of departure, one could use the six datasets of this study, and observe in depth the behavior of the PLS path modeling algorithm.

Despite the limitations, there are several valid conclusions that can be drawn. Firstly and most importantly, this contribution is the first one ever that reports a case of non-convergence of PLS path modeling. I could show that PLS does not always converge. The further search for a proof of convergence, at least for the general PLS path modeling algorithm, can thus be abandoned. Table 2 cumulates the knowledge about the convergence of the PLS path modeling algorithm. Secondly, I found that contrary to common belief, the initialization of the latent variable scores does play a role. In case of all six datasets, convergence could be obtained for some initializations (e.g., average over the indicators), but not for the one used. This has an implication in particular for PLS path modeling software developers. They should choose the average of the indicators as the standard initialization of the latent variable scores. Last but not least, the PLS path modeling algorithm forms part of a wider family of algorithms, the so-called fix-point estimation techniques. It is known that some algorithms of this family do not converge (cf. [Lyttkens 1973](#)). For those algorithms of that family whose convergence has not been proven so far, it may be more promising to look for counterexamples than to try proving convergence.

For users of PLS path modeling, several recommendations have to be spelled out. Practical convergence does not necessarily mean that the PLS algorithm has come to stable estimates. As the second case of non-convergence illustrated, it is well possible

**Table 2** Convergence of the PLS path modeling algorithm

Algorithm	Inner weighting scheme	One or two latent variables		More than two latent variables	
		Mode A	Mode B	Mode A	Mode B
Wold	Centroid scheme	Converges	Converges	Not proven	Converges
	Factorial scheme	Converges	Converges	Not proven	Converges
	Path weighting scheme	Converges	Converges	Not proven	Not proven
Lohmöller	Centroid scheme	Converges	Converges	Not proven	Not proven
	Factorial scheme	Converges	Converges	Does not always converge	Not proven
	Path weighting scheme	Converges	Converges	Does not always converge	Not proven

that the stop criterion may be fulfilled at a certain point, although the algorithm in fact would not converge if it were allowed to continue. Furthermore, users should not specify their path models in a way too similar to the example model. In particular, users should avoid applying the path weighting scheme, estimating incomplete models, and using only one indicator as initial proxy of the latent variable scores. In order to gain confidence in the stability of the PLS estimates, users should triangulate the PLS path modeling results, i.e. estimate the same model with all inner weighting schemes, and compare the results. Finally, it should be noted that the non-convergence of PLS path modeling does not mean that researchers in behavioral, social, and business science should not use PLS any more. Many of the most important psychometric methods, as for instance common factor analysis or covariance-based structural equation modeling, face the issue of non-convergence. It is just that users must learn how to deal with it. Yet, the so far prevailing assumption that PLS always converges has to be withdrawn.

**Acknowledgments** The author thanks Christian M. Ringle and Vincenzo Esposito Vinzi for double-checking the findings by means of experimental versions of their respective PLS path modeling software implementations, SmartPLS and XLSTAT-PLSPM. This paper was initially presented at the 5th International Symposium on Partial Least Squares and Related Methods, Ås, Norway, 5–7 September 2007.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Addinsoft SARL (2007–2008) XLSTAT-PLSPM. Paris, France. <http://www.xlstat.com/en/products/xlstat-plspm/>
- Bagozzi RP, Yi Y (1994) Advanced topics in structural equation models. In: Bagozzi RP (ed) *Advanced methods of marketing research*. Blackwell, Oxford, pp 1–51
- Bass B, Avolio B, Jung D, Berson Y (2003) Predicting unit performance by assessing transformational and transactional leadership. *J Appl Psychol* 88(2):207–218

- Carroll J (1968) A generalization of canonical correlation analysis to three or more sets of variables. *Proceedings 76th Ann Convent Am Psychol Assoc* 3:227–228
- Chin WW (1998) The partial least squares approach to structural equation modeling. In: Marcoulides GA (ed) *Modern methods for business research*. Lawrence Erlbaum Associates, Inc., Mahwah, pp 295–336
- Chin WW, Newsted PR (1999) Structural equation modeling analysis with small samples using partial least squares. In: Hoyle RH (ed) *Statistical strategies for small sample research*. Sage Publications, Thousand Oaks, pp 334–342
- Chin WW, Marcolin BL, Newsted PR (2003) A partial least squares latent variable modeling approach for measuring interaction effects. Results from a monte carlo simulation study and an electronic-mail emotion/adoption study. *Inf Syst Res* 14(2):189–217
- Diamantopoulos A, Winklhofer H (2001) Index construction with formative indicators: an alternative to scale development. *J Market Res* 38(2):269–277
- Dijkstra TK (1981) Latent variables in linear stochastic models: Reflections on “Maximum Likelihood” and “Partial Least Squares” methods. PhD thesis, Groningen University, Groningen, a second edition was published in 1985 by Sociometric Research Foundation
- Dijkstra TK (2009) Latent variables and indices: Herman Wold’s basic design and partial least squares. In: Vinzi VE, Chin WW, Henseler J, Wang H (eds) *Handbook of partial least squares: concepts, methods, and applications, computational statistics, vol II*. Springer, Heidelberg (in print)
- Escofier B, Pagès J (1994) Multiple factor analysis (afmult package). *Comput Stat Data Anal* 18:121–140
- Fornell C (1992) A national customer satisfaction barometer: the swedish experience. *J Market* 56(1):6–21
- Fornell C, Bookstein FL (1982) Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. *J Market Res* 19(4):440–452
- Fornell C, Barclay D, Rhee B (1988) A model and simple iterative algorithm for redundancy analysis. *Multivar Behav Res* 23(3):349–360
- Fu JR (2006) *VisualPLS—Partial Least Square (PLS) Regression—An Enhanced GUI for LVPLS (PLS 1.8 PC) Version 1.04*. National Kaohsiung University of Applied Sciences, Taiwan, ROC. <http://www2.kaus.edu.tw/prof/fred/vpls/index.html>
- Garthwaite P (1994) An interpretation of partial least squares. *J Am Stat Assoc* 89(425):122–127
- Gefen D, Straub D (1997) Gender differences in the perception and use of e-mail: an extension to the technology acceptance model. *MIS Q* 21(4):389–400
- Hanafi M (2007) PLS path modelling: computation of latent variables with the estimation mode B. *Comput Stat* 22(2):275–292
- Hanafi M, Qannari E (2005) An alternative algorithm to the PLS B problem. *Comput Stat Data Anal* 48(1):63–67
- Horst P (1961) Relations among  $m$  sets of measures. *Psychometrika* 26(2):129–149
- Hulland J (1999) Use of partial least squares (pls) in strategic management research: A review of four recent studies. *Strateg Manag J* 20(2):195–204
- Hwang H, Takane Y (2004) Generalized structured component analysis. *Psychometrika* 69(1):81–99
- Israels A (1984) Redundancy analysis for qualitative variables. *Psychometrika* 49(3):331–346
- Krijnen W, Dijkstra T, Gill R (1998) Conditions for factor (in) determinacy in factor analysis. *Psychometrika* 63(4):359–367
- Li Y (2005) *PLS-GUI—Graphic user interface for partial least squares (PLS-PC 1.8)—version 2.0.1 beta*. University of South Carolina, Columbia
- Lohmöller JB (1987) *LVPLS 1.8 Program manual: latent variable path analysis with partial least squares estimation*. Zentralarchiv für Empirische Sozialforschung, Universität zu Köln, Cologne, Germany
- Lohmöller JB (1989) *Latent variable path modeling with partial least squares*. Physica, Heidelberg
- Lytkens E (1973) The fix-point method for estimating interdependent systems with the underlying model specification. *J R Stat Soc Sers A (General)* 136(3):353–394
- R Development Core Team (2007) *R: a language and environment for statistical computing*. R foundation for statistical computing, Vienna, Austria. <http://www.R-project.org>
- Rigdon EE (1998) Structural equation modeling. In: Marcoulides GA (ed) *Modern methods for business research*. Lawrence Erlbaum Associates, Mahwah, pp 251–294
- Ringle CM, Wende S, Will A (2007) *SmartPLS 2.0 M3*. University of Hamburg, Hamburg, Germany. <http://www.smartpls.de>
- Soft Modeling, Inc (1992–2002) *PLS-Graph Version 3.0*. Houston, TX. <http://www.plsgraph.com>
- Tenenhaus M, Vinzi VE (2005) PLS regression, PLS path modeling and generalized Procrustean analysis: a combined approach for multiblock analysis. *J Chemom* 19(3):145–153

- Tenenhaus M, Vinzi VE, Chatelin YM, Lauro C (2005) PLS path modeling. *Comput Stat Data Anal* 48(1):159–205
- Test and Go (2006) SPAD Version 6.0.0. Paris, France
- Ulaga W, Eggert A (2006) Value-based differentiation in business relationships: gaining and sustaining key supplier status. *J Market* 70(1):119–136
- Wold HOA (1966) Non-linear estimation by iterative least squares procedures. In: David FN (ed) *Research papers in statistics*. Wiley, London, pp 411–444
- Wold HOA (1973) Nonlinear iterative partial least squares (NIPALS) modelling. Some current developments. In: Krishnaiah PR (ed) *Proceedings of the 3rd international symposium on multivariate analysis*. Dayton, OH, pp 383–407
- Wold HOA (1974) Causal flows with latent variables: partings of the ways in the light of NIPALS modelling. *Euro Econ Rev* 5(1):67–86
- Wold HOA (1982) Soft modelling: the basic design and some extensions. In: Jöreskog KG, Wold HOA (eds) *Systems under indirect observation. Causality, structure, prediction, vol II*. North-Holland, Amsterdam, pp 1–54
- Wold HOA (1985) Partial least squares. In: Kotz S, Johnson NL (eds) *Encyclopaedia of statistical sciences*, vol 6. Wiley, New York, pp 581–591
- Wold HOA (1985) Partial least squares and LISREL models. In: Nijkamp P, Leitner H, Wrigley N (eds) *Measuring the unmeasurable*. Nijhoff, Dordrecht, pp 220–251
- Wold HOA (1989) Introduction to the second generation of multivariate analysis. In: Wold HOA (ed) *Theoretical empiricism a general rationale for scientific model-building*. Paragon House, New York, pp VIII–XL