SOME REMARKS ON NONLINEAR RELATIONSHIPS IN PLS PATH MODELING

Salvatore Ingrassia[*], Laura Trinchera[**]

[*] *Dipartimento di Impresa Culture e Società, Università di Catania, Corso Italia, 55 - 95129 Catania (Italy), s.ingrassia@unict.it*

[**] *Dipartimento di Matematica, Statistica Università di Napoli "Federico II", Via Cintia, 27 - 80126 Napoli (Italy) ltrinche@unina.it*

Abstract

Structural Equation Models (SEMs) are widely used to model complex causal relationships, especially in economic and social domains. In this framework, PLS Path Modeling algorithm is a widely used technique. However, PLS Path Modeling assumes linear relationships among the latent concepts, and between the observed variables and the corresponding latent variables. Recent studies in marketing research, have pointed out that this hypothesis may appear to be too restrictive. In this paper nonlinear approaches to PLS Path Modeling are discussed and an application to real data concerning the impact analysis of the European Union enlargement to 25 countries on Italian firms is presented.

Keywords Nonlinearity, Partial Least Squares, Structural Equation Models.

1. INTRODUCTION

Structural Equation Models (SEMs) (Bollen, 1989; Kalplan 2000) are widely used for modeling complex causal relationships. The idea behind these models is that the real word complexity can be modeled by using several indicators (the manifest variables, MV) measuring more complex concepts that cannot be directly observed, i.e the latent variables (LVs). Each Structural Equation Model is composed of two different parts: a measurement model and a structural model. The measurement model describes the link between each manifest variable and the corresponding latent variable, while the structural model takes into account the cause-and-effect relations among the latent variables. The parameters of the measurement model are the outer weights (i.e. the weights linking each manifest variable to the correspon-ding latent variable) and the loadings (i.e. the correlation between each manifest variable in the model and the corresponding latent variables); the parameters of the structural model are the path-coefficients, i.e. the regression coefficients linking each endogenous latent variable to its exogenous latent variable(s).

Several approaches exist to estimate such parameters; in particular the PLS-Path Modeling (PLS-PM) algorithm (Wold, 1975; Tenenhaus *et al.*, 2005) is a widely used technique thanks to its features. Namely PLS-PM does not rely on a specific distributional hypothesis. Moreover, according to Tenenhaus (2008) it provides systematic convergence of the algorithm due to its simplicity; it allows data to be managed with a small number of individuals and a large number of variables; it provides a practical interpretation of the latent variable estimates; and it represent a general framework for multi-block analysis.

However, in such models the relations defining both the measurement and the structural models are assumed to be linear. Recent studies in the marketing area have highlighted that such a linear assumption does not fit the asymmetric and the nonlinear nature of these links, see e.g. Anderson and Mittal (2000). In particular, modeling the performance-satisfaction link as symmetric and linear, one can incorrectly estimate the attribute importance weights and therefore misses the mark in prioritizing efforts to maintain and improve customer satisfaction. Modeling an attribute that has an asymmetric and nonlinear relationship with customer satisfaction as symmetric and linear systematically misestimates the impact of that attribute on customer satisfaction.

The rest of the paper is organized as follows. In the next section we introduce the PLS Path Modeling approach to Structural Equation Models and afterwards present a short review of the literature in nonlinear PLS Path Modeling; in section 3 we present an application concerning a case study about impact analysis. Finally in section 4 we give conclusions and perspectives for future works.

2. THE STANDARD PLS PATH MODELING ALGORITHM AND NONLINEAR APPROACHES

PLS Path Modeling (PLS-PM) is an iterative algorithm allowing us to compute Structural Equation Model parameters without making any kind of distributional hypothesis on both manifest and latent variables. Specifically, PLS-PM estimates, through a system of interdependent equations based on simple and multiple regressions, the network of relations among the manifest variables and their own latent variables, and among the latent variables inside the model.

In a PLS Path Model the relations between each manifest variable and the corresponding latent variables is modeled by means of the measurement model. Two different kinds of measurement model schemes can be described according to the type of relations supposed to exist between the manifest variables and the corresponding latent variable: the formative scheme and the reflective scheme. These two schemes assume a different idea of the latent variable. As a matter of fact, in a reflective scheme each manifest variable reflects the corresponding latent variable, thus it is related to the latent variable by a simple regression model. In a formative scheme, instead, each latent variable is obtained as a linear combination of the manifest variables of the block. Thus the measurement model can be expressed as a multiple regression model and the manifest variables are considered as predictors of the corresponding latent variable.

The relations among the several observed blocks of variables are modeled by means of the structural model. Let ξ_j be a generic endogenous latent variable in the model, then structural relationship describing the causations among the latent variables is defined as:

$$\xi_j = \beta\, \Theta_M + \xi_j$$

where Θ_M is the N by M matrix containing the ξ_m latent variables impacting on ξ_j, β is the vector containing the structural parameters linking the M exogenous latent variables to ξ_j , i.e. the path coefficients, and ξ_j is a residual term associated to the generic endogenous latent variable.

In PLS Path Modeling the outer weights (i.e. the weights linking each manifest variable to the corresponding latent variable) and the latent variable scores (ξ_q) are estimated by means of an iterative procedure. Given a set of initial weights, the latent variable scores are obtained through an iterative procedure based on the alternation of the *outer* and the *inner* estimations, till convergence is reached. In particular, in the external estimation, each latent variable is estimated as a linear combination of its own manifest variables; in the internal estimation, each latent variable is estimated by considering its links with the other adjacent latent variables. The internal estimates are then used to update the outer weights. Different approaches are available in order to update the outer weights depending on the kind of measurement model (i.e. the *formative* or the *reflective* scheme), as well as to compute the inner estimates; for further details, see e.g. Tenenhaus *et al.* (2005).

The path coefficients (β_{mj}) linking the *j*-th endogenous latent variable to the *m*-th exogenous latent variable are then estimated using a simple or multiple regression model acting on the estimated latent variable scores.

The above considerations show that all the relationships in the PLS-PM are supposed to be linear. However, this assumption appears to be too restrictive in some applicative cases but, despite practical interest in modeling nonlinear approaches, until now relatively little work has been done in this area. To the authors' knowledge, four approaches exist in modeling nonlinear relationships in PLS Path Modeling framework, as elaborated by Wold (1982), Chin *et al.* (2003) and Henseler *et al.* (2008), Krämer (2005) and finally, by Jakobowicz (2007).

The easiest way to consider nonlinearity among latent variables in a PLS-PM is to add interaction effects to the model. Chin *et al.* (2003) firstly introduced interaction effects in PLS-PM. The idea is to model the interaction terms by means of the product-indicators of the centered exogenous manifest variables, see Figure 1. Of course, the number of indicators for the latent interaction term is equal to the product between the number of manifest variables in each exogenous block and thus it easy to notice that this directly increases the number of manifest variables in the model. Moreover, we remark that the latent interaction term obtained by its indicators does not coincide with the product of the exogenous latent variable scores.

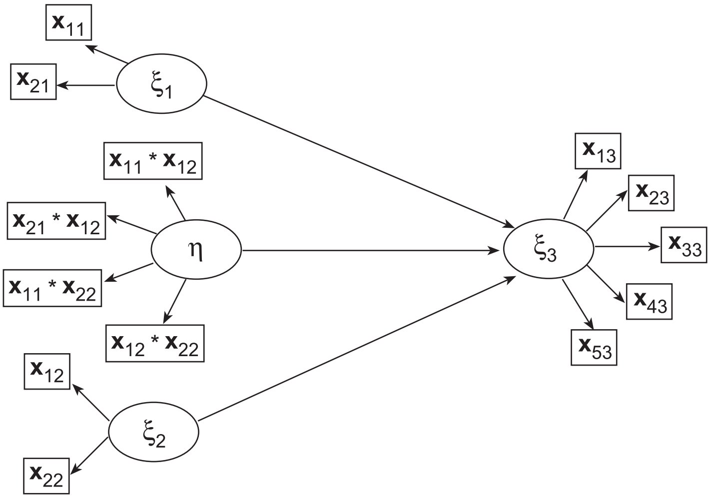

Fig. 1: Interaction effect modeled by means of product of indicators.

To overcome these problems, Henseler *et al.* (2008) recently proposed a two-stage procedure to model interaction effects, see Figure 2. In particular, in a first stage a standard PLS-PM analysis is run in order to estimate the direct effect of each exogenous latent variable as well as the latent variable scores. In a second stage, the interaction effect is obtained as the product of endogenous latent variables scores. Then, the exogenous latent variable scores (estimated in the first-stage), and the interaction effect (obtained in the second-stage) are used as exogenous variables in a multiple linear regression to obtain path coefficients. We remark that here the nonlinearity is modeled just by introducing an interaction term in the structural model. In our opinion this approach is not satisfactory.

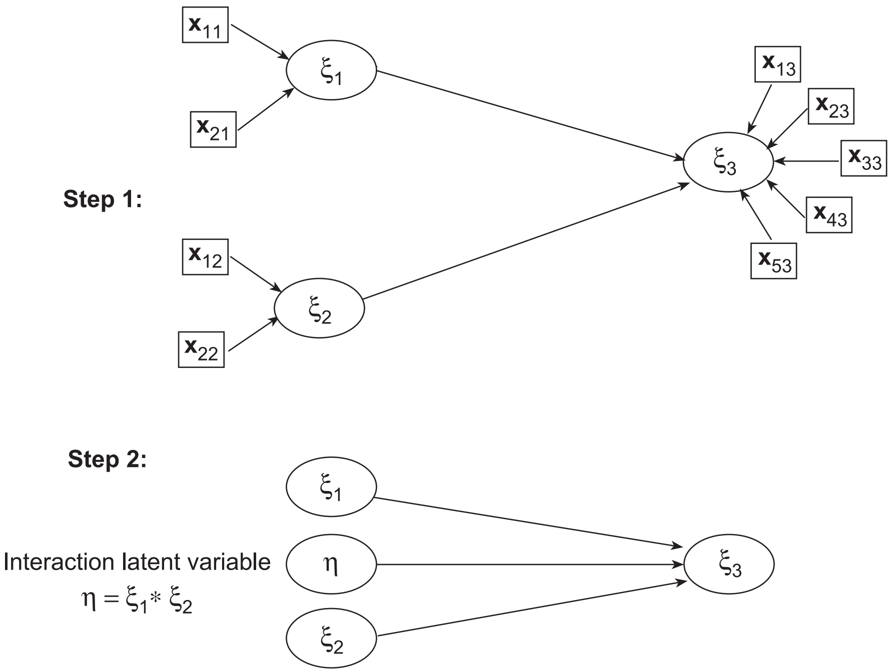

Fig. 2: Interaction effect modeled by means of product of indicators.

A modification in the iterative algorithm originally proposed about twenty years previously by Wold (1982) can be considered as a compromise between the two-stage approach of Henseler (2008) and that of Chin *et al.* (2003). Wold (1982) suggested modeling nonlinearity in the structural model by adding a step to the classical PLS-PM algorithm. In particular, once the outer estimation of each latent variable has been computed, Wold proposed computing interaction term proxies as an element-wise product of the outer estimates of the latent variables. Then, the structural model is augmented in order to obtain inner estimates of endogenous latent variables by considering also the interaction term proxies. The inner estimates of the latent variables are then used to update the outer weights as usual in the PLS-PM algorithm. Unlike the two-stage procedure of Henseler *et al.* (2008), Wold's approach allows the updating of the interaction term while the algorithm has been carried out. Moreover, unlike the method of Chin *et al.* (2003), in Wold's approach the interaction term is modeled only in the structural model and this allows consideration of the problem of obtaining indicators for the interaction terms. Anyway, this approach does not take into account nonlinear links between the manifest and latent variables.

More recently, Krämer (2005) proposed a new approach for handling nonlinearity in PLS-PM. Here the basic hypothesis is that the nonlinearity affects the outer model. In other words, nonlinear relations are supposed to exist between each manifest variable and the corresponding latent variable; the nonlinear measurement relations are modeled by means of a suitable data transformation called *kernel trick*. Indeed, following this approach each relation in the measurement model can be regarded as a simple/multiple regression problem according to the chosen outer model scheme; in particular this approach has been developed in a formative scheme and thus each relation in the measurement model can be approximated by means of some function (not necessarily a linear one). This amounts to considering a kernel transformation on the block of manifest variables and then applying a standard PLS-PM algorithm on such transformed data. Here the choice of a suitable kernel function can be considered as an additional parameter in the model. In general, a kernel function should be identified for each block in the model but, for simplicity, Krämer (2005) suggested using only one family of kernels for all the blocks in the model. Furthermore, the same kernel parameters could be used across the blocks. However, such an approach does not provide any estimation of the outer weights and thus the interpretation of the link between each manifest variable and the corresponding latent variables is no longer possible. This represents an important lack of interpretability in a model aiming at explaining the relationships between latent and manifest variables.

Finally, recently Jakobowicz (2007) proposed modifying PLS-PM algorithm by adding a new step in order to transform some latent variables in an optimal way. Based on an idea of Coolen *et al.* (1981), who combined optimal scaling techniques with B-splines, Jakobowicz (2007) applied a B-spline transformation to some latent variables in the model. However, this approach requires identifying a well-established target latent variable in the model. In particular, once the target latent variable is chosen, a B-spline transformation is applied to each exogenous latent variable impacting on the target latent variable. In complex models the identification of a well-established target latent variable is not immediate, making it hard to carry out such a procedure. It is important to remark that both Krämer (2005) and Jakobowicz (2007) pointed out that model quality did not improve by taking into account nonlinear transformations in PLS-PM. As a matter of fact, because SEM's are complex models *per se,* it does not seem suitable to introduce nonlinearities into the model in a direct way.

These considerations suggested another approach in order to tackle our problem. The idea is to approximate nonlinearity using piece-wise linear sub-models. This approach is appealing because it is simple and not very computationally demanding. In PLS-PM literature several techniques have been proposed to obtain local models from a unique global model (Hahn *et al.*, 2002 and Ringle *et al.*, 2008; Sanchez and Aluja, 2007). In this paper we shall consider the REBUS-PLS (REsponse Based Unit Segmentation in PLS Path Model) which is an iterative algorithm for detecting unobserved heterogeneity in PLS Path Models, see Trinchera (2007) and Esposito Vinzi *et al.* (2008). This approach is a very interesting approach because it allows us to estimate at the same time both the unit memberships to latent classes and the class specific parameters of the local models without making any kind of distributional assumption either on the manifest variables or on the latent variables. The core of the algorithm is a so-called Closeness Measure (CM) between units and models based on residuals. The idea behind the definition of this new measure is that if latent classes exist, units belonging to the same latent class will have similar local models, i.e. similar performance as regards the global model. Moreover, if a unit is assigned to the correct latent class, its performance in the local model computed for that specific class will be better than the performance obtained by the same unit considered as supplementary in all the other local models. The CM measure is based on two kinds of residuals: the measurement residuals (i.e., the residuals of the regressions of each manifest variable over it own latent variable) and the structural residuals (i.e., the residuals of the regressions of the endogenous latent variables over their respective explanatory latent variables). The measurement residuals are computed in order to take into account units' performances in the measurement model. While, the structural residuals considers the units' performances in the structural model. This allows us to obtain unit clustering according to both the measurement and structural models. In each model we have P measurement residuals, one for each manifest variable in the model, and J structural residuals, corresponding to the number of endogenous latent variables. For a thorough description of the REBUS-PLS algorithm and the computation of the measurement and the structural residuals, please refer to the original REBUS-PLS papers (Trinchera, 2007; Esposito Vinzi et al., 2008). It is important to notice that both the measurement and the structural residuals are computed for each unit as respect to each local model regardless the memberships of the unit to the specific latent class. To conclude, once the stability on class composition is reached, final local models are computed. The class-specific parameters are then compared in order to explain differences among detected latent classes.

3. A CASE STUDY ON IMPACT ANALYSIS

The proposed method will be applied to a case-study referring to the impact of the European enlargement to 25 members on Italian firms. Such a case has already been studied using linear methods (see Lauro and Scepi, 2008). The use of nonlinear methods for the analysis of causal relations may lead to develop more effective models and to enhance the economic understanding of the underlying phenomena.

This analysis involves only N=158 units collected by Lauro and Scepi (2008). The data come from a survey conducted on big and medium Italian manufacturing industries. The questionnaire was based on 22 items, divided into 6 sections, see Table 1. Each indicator is measured on a 0-10 scale. In this paper we do not provide an extensive presentation neither of the survey data nor of the form of the model; see Lauro and Scepi (2008) for details about these issues. Here we are interested in evaluating if modeling nonlinear relations by means of piece-wise linear sub-models can improve model quality, giving more accurate interpretation of the coefficients.

The path model for the evaluation of the impact of the EU enlargement is composed of 6 latent variables. The idea behind the model is that the perceived impact of the EU enlargement (IMPACT) depends on the following latent concepts: the perceived direct threats (THREATS), the opportunities (OPPORTUNITIES), the scenario effects (SCENARIO) and the decisions of the Public Administration (PUBLIC ADMIN.). The feedback (FEEDBACK) of the enterprises depends on the perceived direct threats (THREATS), on the opportunities (OPPORTUNITIES) and on the perceived impact (IMPACT). Moreover, the THREATS impact on the PUBLIC ADMINISTRATION, which, in turn, impacts on the SCENARIO effects. Finally, the OPPORTUNITIES depend on the SCENARIO. All indicators are supposed to be linked to the corresponding latent variable through a reflective measurement model, see Figure 3.

First a standard PLS-PM analysis on all the 158 units has been performed through XLSTAT 2008 software. The results of the global model are presented in Figure 4 and in Tables 2, 3. The structural model results show that there are least two non-significant relations in the structural model according to bootstrap intervals and p-values. Namely, the latent variable OPPORTUNITIES has no impact on the FEEDBACK capability of the enterprises, with a p-value of 0.92 associated to a path coefficient of 0.01. And the variable THREATS does not influence the variable IMPACT, because of a p-value of 0.74 associated to a path coefficient of 0.02. Moreover, another two relations (THREATS on FEEDBACK, and PUBLIC ADMINISTRATION on IMPACT) can be considered significant because of a p-value less than 0.1. Among the exogenous latent variables impacting on the endogenous latent variable IMPACT, the latent variable OPPORTUNITIES exhibits a strong relationship, having a path coefficient equal to 0.45. The arrows thickness in Figure 4 depends on the statistical significance of the relations. The higher the statistical significance of the path coefficient associated to the relation, the thicker the arrow is. Table 2 provides a complete list of the structural model results.

Tab. 1: Measurement instruments for the impact analysis on Italian enterprises.

Latent variables	***Manifest Variables***
Perceived direct opportunities	investments: investment profitability (productive and financial) in the new EU members; market: Potential market expansion for Italian enterprises; human resources: labour costs reduction for a wider access to skilled human resources under competitive conditions.
Perceived direct threats	investments: Investments reduction (productive and financial) in Italy (outflow of capital and resources towards new EU partners); market: worsening of internal market reduction; human resources: contraction of the inner size due to the increase of domestic unemployment and labour cost reduction.
Scenario effects	security: greater security as a result of coordination among national policies on justice, organized crime, terrorism; cooperation: increase of cooperation in society (especially regarding migration) and the environment; reputation: increase in political reputation on international matters; normative unification: homogeneization of the legal systems (production and competition) and of the standards for quality certification; customs union: beneficial on European trade as a result of the establishment of a customs union.
Expectations on Public Administration	information: accessibility and transparency of information on the economies of the new members and on EU legislation; training: promotion of training initiatives and professional improvement to increase EU trade; regulation: adoption of measures to ensure the complete fulfilment of EU initiatives and fair competition; funding: funds to protect national business from external competition, especially backward and border areas.
Perceived impact on the industry	value of the stock: rise in the business and in the overall value of the industry; stakeholders benefits: larger profits and advantages for the stakeholders; trust in Europe: greater competitive capability of the system in the world; size of the firm: rising trend in the size and in the globalization of the firms in the industry.
Expected business feedback	human resources: supply of human resources in the new members; organization: reorganization of business processes; strategy: exploitation of new distribution channels and development of new markets and/or of parts of them.

Tab. 2: Structural model results for the global model. Bootstrap confidence Linterval: * *p*-value less than 0.1; ** *p*-value less than 0.05; * *p*-value less than 0.01.**

Endogenous LV	*Exogenous LVs*	*Path Coeff.*	*Lower bound (95%)*	*Upper bound(95%)*
Opportunities	R^2	0.26		
	Scenario	0.51^{***}	0.29	0.70
Scenario	R^2	0.12		
	Public Administration	0.34^{***}	0.15	0.56
Public Administration	R^2	0.09		
	Threats	-0.30^{***}	-0.49	-0.16
Impact	R^2	0.44		
	Threats	0.02	-0.16	0.24
	Public Administration	0.16^{**}	-0.03	0.34
	Scenario	0.23^{***}	0.07	0.40
	Opportunities	0.45^{***}	0.23	0.61
Feedback	R^2	0.32		
	Threats	-0.13^{*}	-0.27	0.02
	Opportunities	0.01	-0.17	0.22
	Impact	0.52^{***}	0.33	0.69

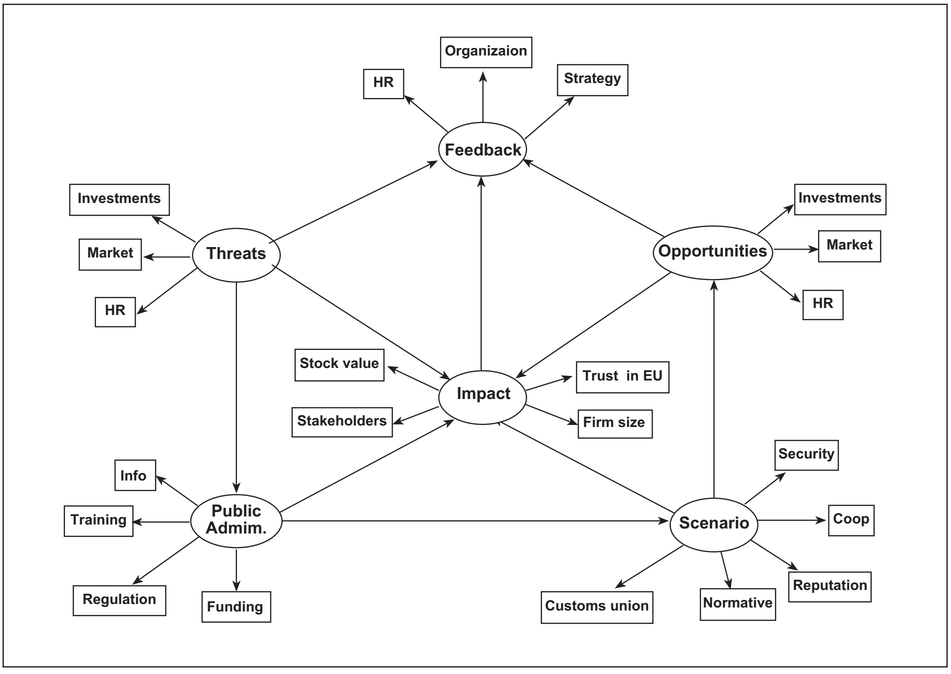

Fig. 3: The path model to evaluate the impact of the EU enlargement to 25 members.

As far as the results of the measurement model are concerned, we notice that the relationships in each block, except for the one linking the Human Resource to the THREATS latent variable, have a factor loading at sufficiently high levels. Table 3 provides an overview of the measurement model results obtained for the global model. In Figure 4 the thickness of the arrows of the measurement model depends on the correlation between each manifest variable and the corresponding latent variable. The higher the correlation between the manifest and the latent variable is, the thicker the arrow is. Moreover, the values on arrows of the measurement model are the normalized weights associated to each manifest variable. In PLS Path Modeling, the global model quality is evaluated by means of the Goodness of Fit Index (GoF), where a high value of the GoF index entails a good model. In our case, the GoF index gives the value 0.386, which is quite modest. The presence of nonlinear relations may cause this small value of the GoF index, as well as the small values of the R^2 in the structural model.

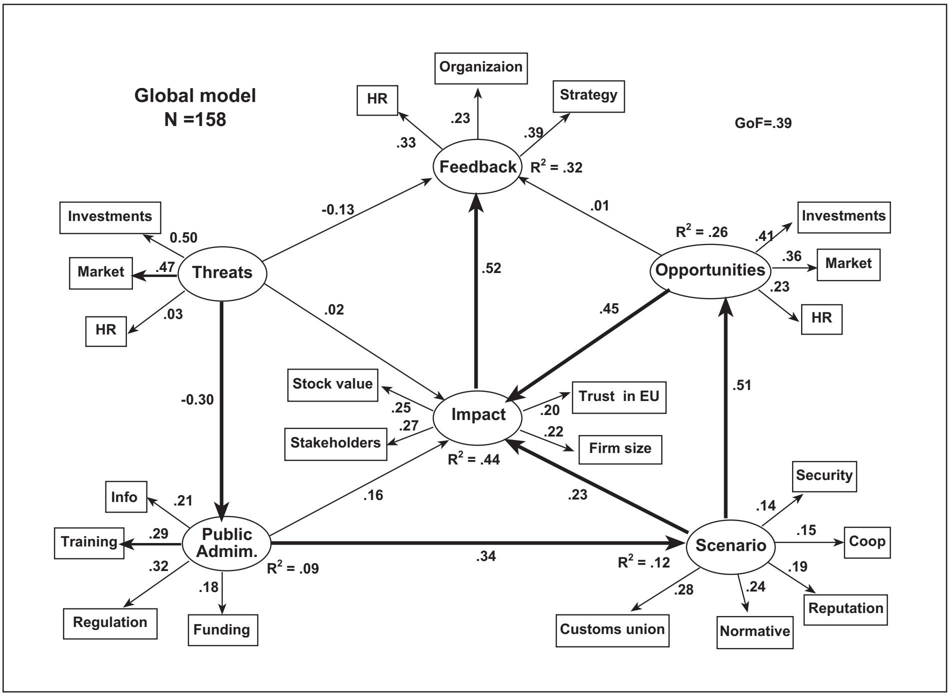

Fig. 4: Global model results for the EU enlargement path model. The thickness of the structural model arrows depends on the statistical significance of the relations. The thickness of the measurement model arrows depends on the correlation between each manifest variable and the corresponding latent variable.

Tab. 3: Measurement model results for the global model.

LV	*MVs*	*Loadings*	*Lower bound*	*Upper bound (95%)*	*Communality (95%)*
Opportunities	D.G. ρ	0.84			
	Investments	0.85	0.76	0.90	0.72
	Market	0.83	0.73	0.90	0.69
	HR	0.56	0.30	0.75	0.31
Threats	D.G. ρ	0.82			
	Investments	0.87	0.20	0.96	0.76
	Market	0.84	0.58	0.98	0.71
	HR	0.45	-0.66	0.71	0.21
Scenario	D.G. ρ	0.86			
	Security	0.72	0.57	0.79	0.52
	Cooperation	0.74	0.57	0.84	0.55
	Reputation	0.68	0.50	0.81	0.47
	Normative	0.78	0.64	0.84	0.61
	Customs union	0.73	0.61	0.80	0.54
Public Administration	D.G. ρ	0.87			
	Information	0.75	0.47	0.90	0.57
	Training	0.88	0.76	0.92	0.77
	Regulation	0.89	0.81	0.93	0.78
	Funding	0.61	0.27	0.81	0.37
Impact	D.G. ρ	0.92			
	Stock	0.88	0.81	0.92	0.77
	Stakeholders	0.91	0.88	0.94	0.82
	Trust	0.85	0.76	0.91	0.73
	Firm size	0.80	0.71	0.86	0.65
Feedback	D.G. ρ	0.84			
	HR	0.79	0.67	0.86	0.63
	Organization	0.78	0.65	0.87	0.61
	Strategy	0.81	0.71	0.88	0.66

The application of REBUS-PLS to the EU enlargement data has led to the identification of two latent classes. The first latent class is composed of 61 units, i.e. close to 39% of the whole sample and thus the remaining units belong to the second latent class. REBUS-PLS analysis has been performed through an R-Package (http://cran.r-project.org/web/packages/plspm/index.html) developed by Sanchez and Trinchera (2009). The local model results are summarized in Figures 5 and 6. Also in these figures, the thickness of the arrows depends on the statistical significance of the path coefficients in the measurement model, and on the correlation between manifest and latent variables in the measurement model. At a first look it is easy to notice that the Figures 5 and 6 are very different as regards both

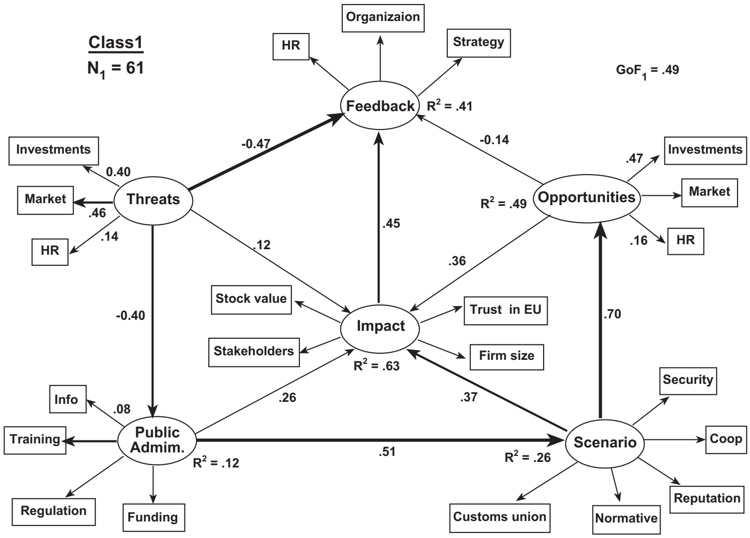

Fig. 5: Local model results for the EU enlargement path model computed for latent class 1. The thickness of the structural model arrows depends on the statistical significance of the relations. The thickness of the measurement model arrows depends on the correlation between each manifest variable and the corresponding latent variable.

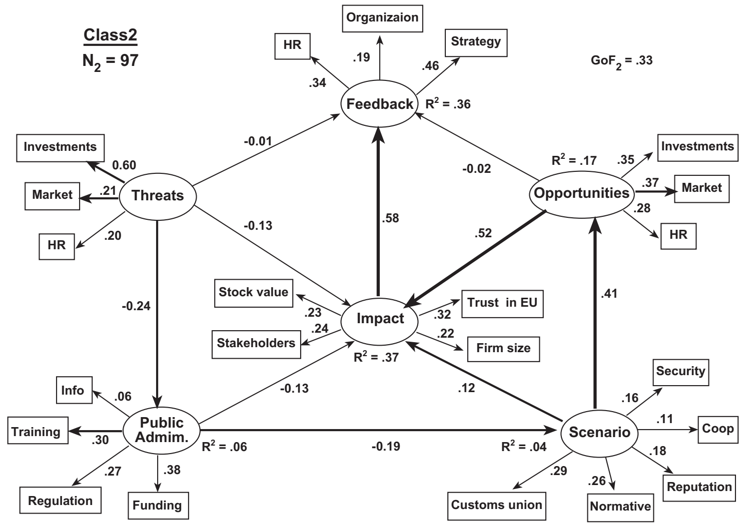

Fig. 6: Local model results for the EU enlargement path model computed for latent class 2. The thickness of the structural model arrows depends on the statistical significance of the relations. The thickness of the measurement model arrows depends on the correlation between each manifest variable and the corresponding latent variable.

the measurement and the structural model results. In particular, Tables 4 and 5 show a first difference in the relations impacting on the latent variable FEEDBACK. As a matter of fact, the only exogenous latent variable with a significant impact for the latent class 2 is the latent variable IMPACT, with a path coefficient of 0.58. On the contrary, in latent class 1 both the latent variables IMPACT and THREATS have a significant impact on the FEEDBACK, having a *p*-value less than 0.005. Moreover, as regards the impacts on the endogenous latent variable IMPACT, it is possible to notice that the latent variable THREATS has no significant impact in both the latent class models, as well as in the global model. Differences arise in the significance of the other relations impacting on the endogenous latent variable IMPACT. In fact, in latent class 2 the only significant path coefficient is the one linking the OPPORTUNITIES to the IMPACT, while in latent class 1, the path coefficients associated to the latent variables OPPORTUNITIES, SCENARIO and PUBLIC ADMINISTRATION are significant (*p*-value less than 0.01 for the first two ones, *p*-value less than 0.011 for the last one). The last interesting difference is focused on the path coefficient linking the exogenous latent variable PUBLIC ADMINISTRATION to the endogenous latent variable SCENARIO. As a matter of fact, this coefficient shows a negative value for latent class 2 (*p*-value less than 0.1), while it is relevant in both the local model computed for latent class 1 and the global model (*p*-value less than 0.01 for both models). As regards measurement model results, the two classes differ in the intensities of the outer weights for the manifest variables in the block THREATS and in the block PUBLIC ADMINISTRATION. Concerning the block of the latent variable THREATS, the external weight associated to the manifest variable Human Resources is definitely higher for latent class 2 than for latent class 1 (see Figures 5 and 6). Finally, concerning the block of the latent variable PUBLIC ADMINISTRATION, we notice that the manifest variable *Info* has no impact in building the latent variable score in both the local models, while in the global model its weight in computing the latent variable score is close to 0.20. A complete overview of the measurement model results for local models computed for latent class one and two is provided respectively in Tables 6 and 7. Here the loadings associated with each manifest variable are presented.

Tab. 4: Structural model results for the local model computed for latent class 1: * *p*-value less than 0.1; ** *p*-value less than 0.05; * *p*-value less than 0.01.**

Endogenous LV	*Exogenous LVs*	*Path Coeff.*	*Lower bound (95%)*	*Upper bound (95%)*
Opportunities	R^2	0.49		
	Scenario	0.70***	0.49	0.86
Scenario	R^2	0.26		
	Public Administration	0.51***	0.20	0.80
Public Administration	R^2	0.12		
	Threats	-0.35***	-0.74	0.18
Impact	R^2	0.63		
	Threats	0.12	-0.25	0.43
	Public Administration	0.26**	0.01	0.47
	Scenario	0.37***	0.15	0.56
	Opportunities	0.36***	0.09	0.59
Feedback	R^2	0.41		
	Threats	-0.47***	-0.68	-0.20
	Opportunities	-0.14	-0.40	0.34
	Impact	0.45***	0.11	0.74

Tab. 5: Structural model results for the local model computed for latent class 2: * *p*-value less than 0.1; ** *p*-value less than 0.05; * *p*-value less than 0.01.**

Endogenous LV	*Exogenous LVs*	*Path Coeff.*	*Lower bound (95%)*	*Upper bound (95%)*
Opportunities	R^2	0.17		
	Scenario	0.41***	0.07	0.62
Scenario	R^2	0.04		
	Public Administration	-0.19*	-0.45	0.31
Public Administration	R^2	0.06		
	Threats	-0.24**	-0.61	0.11
Impact	R^2	0.37		
	Threats	-0.13	-0.34	0.29
	Public Administration	-0.13	-0.34	0.21
	Scenario	0.12	-0.07	0.40
	Opportunities	0.52***	0.28	0.72
Feedback	R^2	0.36		
	Threats	-0.01	-0.26	0.17
	Opportunities	0.02	-0.21	0.31
	Impact	0.58***	0.34	0.80

Tab. 6: Measurement model results for the local model computed for latent class 1.

LV	*MVs*	*Loadings*	*Lower bound (95%)*	*Upper bound (95%)*	*Communality*
Opportunities	D.G. ρ	0.82			
	Investments	0.88	0.81	0.94	0.78
	Market	0.86	0.67	0.94	0.74
	HR	0.51	-0.11	0.76	0.26
Threats	D.G. ρ	0.86			
	Investments	0.84	0.33	0.92	0.71
	Market	0.89	0.65	0.95	0.79
	HR	0.66	0.12	0.88	0.44
Scenario	D.G. ρ	0.88			
	Security	0.76	0.53	0.86	0.58
	Cooperation	0.80	0.64	0.89	0.64
	Reputation	0.82	0.60	0.91	0.67
	Normative	0.76	0.47	0.87	0.57
	Customs union	0.72	0.47	0.86	0.51
Public Administration	D.G. ρ	0.71			
	Information	0.32	-0.59	0.88	0.11
	Training	0.83	0.47	0.93	0.70
	Regulation	0.85	0.12	0.93	0.72
	Funding	0.83	0.14	0.93	0.70
Impact	D.G. ρ	0.93			
	Stock	0.91	0.85	0.94	0.82
	Stakeholders	0.93	0.88	0.96	0.87
	Trust	0.87	0.76	0.92	0.75
	Firm size	0.77	0.49	0.91	0.60
Feedback	D.G. ρ	0.87			
	HR	0.83	0.73	0.92	0.69
	Organization	0.87	0.81	0.93	0.76
	Strategy	0.78	0.49	0.92	0.60

Tab. 7: Measurement model results for the local model computed for latent class 2.

LV	***MVs***	***Loadings***	***Lower bound (95%)***	***Upper bound (95%)***	***Communality***
Opportunities	D.G. ρ	0.79			
	Investments	0.81	0.63	0.89	0.66
	Market	0.82	0.60	0.92	0.67
	HR	0.60	0.27	0.80	0.36
Threats	D.G. ρ	0.80			
	Investments	0.87	0.20	0.96	0.76
	Market	0.84	0.58	0.98	0.71
	HR	0.45	-0.66	0.71	0.21
Scenario	D.G. ρ	0.83			
	Security	0.69	0.20	0.84	0.48
	Cooperation	0.68	0.30	0.85	0.46
	Reputation	0.58	0.15	0.83	0.34
	Normative	0.78	0.30	0.88	0.60
	Customs union	0.70	0.38	0.85	0.49
Public	D.G. ρ	n/a			
Administration	Information	0.38	-0.01	0.83	0.15
	Training	0.77	0.15	0.91	0.59
	Regulation	0.64	-0.01	0.88	0.41
	Funding	0.66	-0.48	0.93	0.44
Impact	D.G. ρ	0.91			
	Stock	0.82	0.74	0.92	0.68
	Stakeholders	0.87	0.80	0.93	0.75
	Trust	0.86	0.77	0.91	0.75
	Firm size	0.82	0.68	0.89	0.67
Feedback	D.G. ρ		0.83		
	HR	0.77	0.59	0.88	0.60
	Organization	0.68	0.30	0.87	0.46
	Strategy	0.85	0.73	0.94	0.73

4. CONCLUSIONS AND PERSPECTIVES

In this paper we have addressed some issues regarding nonlinear PLS Path Modeling. First, on the basis of some theoretical justifications, we have highlighted that such problems arise in important areas like marketing and economy. Secondly, we have remarked that poor research has been done in this field, as it is illustrated by the modest number of papers devoted to nonlinear PLS-PM. The proposals presented in literature can be grouped in four different approaches, even if in two of them (Wold, 1982; Chin *et al.*, 2003 and Henseler *et al.*, 2008) the nonlinearity is taken into account only by adding some interaction terms and then they cannot be considered true nonlinear approaches at all. The other two approaches present respectively lack of interpretability (Krämer, 2005) and difficulties in practical implementation of the procedure (Jakobowicz, 2007). Moreover, both Krämer (2005) and Jakobowicz (2007) remarked that model quality did not improve by taking into account nonlinear transformations in PLS-PM.

For this reason, in this paper we have proposed another approach in order to model nonlinearities, i.e. we suggest approximating nonlinearity using piece-wise linear sub-models, based on the REBUS-PLS approach which aims at obtaining local models that fit better than the global model.

Such ideas have been applied in a case study concerning the European enlargement to 25 members on Italian firms. The application of REBUS-PLS led to the identification of two latent classes and gave a deeper comprehension of the relationships among the variables. Actually some negligible path coefficients (in the unique global model) come from the averages of opposite values in the two different sub-models. However, more investigations are necessary and this provides material for further research. Finally, we remark that REBUS-PLS concerns models showing a reflective measurement model. Developments of the REBUS-PLS algorithm to take into account formative indicators as well are ongoing.

REFERENCES

ANDERSON E.W. and MITTAL V. (2000), Strengtening the satisfaction-profit chain, *Journal of Service Research*, **3**, 107-120.

BOLLEN K. A. (1989), *Structural equations with latent variables*, Wiley, New York.

CHIN W., MARCOLIN B.L., and NEWSTED P.N. (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, *Information Systems Research*, **14**, 189-217.

COOLEN H., VAN RIJCKEVORSEL J., and DE LEEUW J. (1982), An Algorithm for Nonlinear Principal Components Analysis with B-Splines by means of Alternating Least Squares, in *COMPSTAT 1982*, Part 2, Physica-Verlag, Wien, 61-63.

ESPOSITO VINZI V. and LAURO C. (2003), PLS regression and classification, in *Proceedings of the PLS'03 International Symposium*, DECISIA, France, 45-56.

ESPOSITO VINZI V., TRINCHERA L., SQUILLACCIOTTI S. and TENENHAUS M. (2008), REBUS-PLS: A response - based procedure for detecting unit segments in PLS path modeling, *Applied Stochastic Models in Business and Industry*, **24**, 439-458.

HAHN C., JOHNSON M., HERMANN A. and HUBER F. (2002), Capturing customer heterogeneity using a finite mixture PLS approach, *Schmalenbach Business Review*, **54**, 243-269.

HENSELER J. and FASSOTT G. (2008), Testing Moderating Effects in PLS Path Models: An Illustration of Available Procedures, in "*Handbook of Partial Least Squares - Concepts, Methods and Applications*, Esposito Vinzi V., Chin W., Henseler J. and Wang H. (Eds)", Springer, Berlin.

JAKOBOWICZ, E. (2007), *Contributions aux modèles d'équations structurelles a variable latentes*, PhD thesis, CNAM, Paris, France.

KAPLAN D. (2000), *Structural Equation Modeling: Foundations and Extensions*, Sage Publications Inc., Thousands Oaks, California.

KRÄMER N. (2005), Nonlinear partial least squares path models, *IASC, 3rd International Conference on Computational Statistics & Data Analysis*.

LAURO C. and SCEPI G. (2008), A latent variable based model for the impact evaluation of the enlargement of the European Union, *Statistica Applicata*, **20** (1), 33-48.

R Development Core Team (2009), *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

RINGLE C., WENDE S. and WILL A. (2008), Finite mixture partial least squares analysis: Methodology and numerical examples, in "*Handbook of Partial Least Squares - Concepts, Methods and Applications*, Esposito VINZI V., CHIN W., HENSELER J. and WANG H. (Eds)", Springer, Berlin.

SANCHEZ G. and ALUJA T. (2007), A simulation study of PATHMOX (PLS Path Modeling segmentation tree) sensitivity, in *PLS'07: 5th International Symposium on PLS and Related Methods*, Oslo, Norway.

SANCHEZ G. and TRINCHERA L. (2009), *plspm: Partial Least Squares data analysis methods*. R package version 0.1-4.

TENENHAUS M. (1998), *La Régression PLS: théorie et pratique*, Technip, Paris.

TENENHAUS M. (2008), Component-based Structural Equation Modelling, *Total Quality Management & Business Excellence*, **19** (7), 871-886.

TENENHAUS M., ESPOSITO VINZI V., CHATELIN Y.M. and LAURO C. (2005), PLS path modeling, *Computational Statistics and Data Analysis*, **48**, 159-205.

TRINCHERA L. (2007), *Unobserved Heterogeneity in Structural Equation Models: a new approach to latent class detection in PLS Path Modeling*, PhD thesis, Department of Mathematics and Statistics, University of Naples "Federico II", Italy.

WOLD H. (1975), Modelling in complex situations with soft information, in *Third World Congress of Econometric Society*, Toronto, Canada.

WOLD H. (1982), Soft Modeling. The Basic Design and Some Extensions, in "*Systems under Indirect Observation. Causality, Structure, Prediction*, Part I, G. Jöreskog and H. Wold (Eds)", North-Holland, Amsterdam.

ALCUNE NOTE SU MODELLI PLS PATH MODELING CON RELAZIONI NON LINEARI

I modelli ad equazioni strutturali sono ampiamente utilizzati per modellare relazioni causali complesse, specie in campo economico e sociale. In questo contesto, l'algoritmo PLS-Path Modeling costituisce una tecnica utilizzata da numerosi autori. Usualmente questo metodo assume relazioni di tipo lineare fra le variabili latenti, e fra le variabili osservate e le corrispondenti variabili latenti. Recenti studi nel settore del marketing hanno evidenziato come tale assunzione possa risultare troppo restrittiva. In questo lavoro vengono discussi approcci non lineari al PLS Path Modeling. Viene inoltre presentata un'applicazione concernente l'analisi dell'impatto dell'allargamento dell'Unione'Europea a 25 membri sulle imprese italiane.