G.Vittadini - P.G. Lovaglio

The estimate of latent variables in a structural model : an alternative approach to PLS

Abstract
In this paper we consider the problems arising from PLS in the context of a recursive structural model with latent variables (LV). The estimates of LV's are performed in such a way that, in some situations, the role of the manifest variables (MV) in the measurement models is not always respected: this inconsistency justifies another definition of LV and a new measurement model by making rational use of the paths between LV, and those linking the indicators of different LV.
The model for the estimate of only LV will be extended to more LV's, and finally, using an RCDR analysis, the parameters matrixes performed satisfy all the properties of the LV in the structural model.

Keyword: Partial Least Squares, Latent Variable, Restricted Component Decomposition Regression. .

1. Interpretation of Latent Variable

LISREL ([6], [8]) is the most used model for structural relations with LV but its drawbacks (non uniqueness of LV scores, normality distribution of LV, strong error hypotheses, lack of sufficient conditions for model identification) has already been shown in depth in previous works ([4],[5], [9], [11], [14], [15], [16]).
PLS ([13], [17]) is a model, proposed in alternative to the LISREL model, based on soft hypoteses. PLS provides the estimate of structural parameters in a second stage, by using the scores of the LV achieved in the first stage as Wold says "by deliberate approximation as linear aggregation (proxy) of its manifest indicators". (For this definition of LV in literature see also Regression Component Analysis, RCD [12]).
In this way, the structural equations become a Path Analysis between two sets of estimated linear combinations.
PLS estimates the scores of two vectors of LV ξ, η, (for the sake of

simplicity unidimensional) specified in the path diagram of Fig.1:

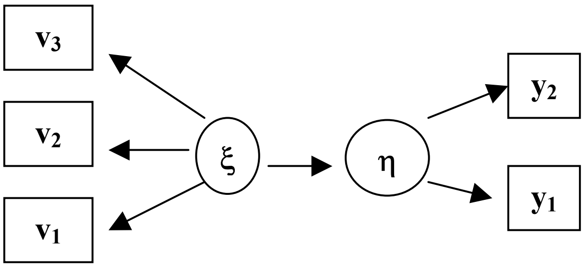

Fig. 1: Path diagram of a structural model, with η and ξ unidimensional.

PLS defines an LV as a weighted combination of its manifest variables, but, as it will be shown later, the weights (and the scores of LV) are performed with arbitrariness by means of regressions which do not respect the role of the manifest variables in the structural model.

2. An example of logic inconsistency of PLS estimates

In a previous paper ([2]) the authors estimated the last equation of a recursive model (Generating Function of Income, GFI, [1])

$$\mathbf{y} = g\,(\mathbf{h}, \mathbf{k}) \tag{1}$$

where the income (**y**) is the dependent variable and the net wealth (**k**) and human capital (**h**) its regressors: the scores of **h** were previously estimated with PLS as the first principal component of a set of manifest variables (income, years of schooling, net wealth, total debt).
Dagum and Vittadini recognized the logical inconsistency of the model specified because of the twofold role of **y**: in the PLS step, **y** (together with other MV's) defines the scores of **h**, whereas in (1) the LV **h** is a predictor of **y**.

3. Logic inconsistency of the manifest indicators in PLS model

a) In a model with only an LV the estimate of η as a linear aggregation of its manifest indicators $\mathbf{y_i}$ does not allow the specification of a linear model where η is a regressor of any $\mathbf{y_i}$ endogenous variable, because it results in unit R^2.

b) in complex models with more than one LV (ξ, η with the corresponding MV's collected in the matrixes **V** and **Y** respectively) PLS performs the weights (**a,c**) that define the scores of LV's $\eta = \mathbf{Ya}$, $\xi = \mathbf{Vc}$ iterating the two following steps:

1) the initial estimate of ξ, η (ξ^o, η^o) as principal components of its indicators;

2) the coefficients (**a, c**) are performed in 4 different ways: *Mode* (A,A) consist of (simple) regression of ξ^o on $\mathbf{y_i}$ and η^o on $\mathbf{x_i}$, *Mode* (B,B) in (multiple) regression of ξ^o on **y** and η^o on **v**, *Mode* (A,B) and *Mode* (B,A) by mixing the previous options.

This methodology is not consistent ([9]) with the factorial equations for ξ ed η specified in Fig.1 because:

- the initial estimates of η and ξ (η_0 ξ_0) are performed by treating the manifest variables as *causes*, whereas these are *indicators* (effects) of its LV; Mode (B,B) updates the estimates of ξ regressing ξ_0 on ($\mathbf{y_1}$ $\mathbf{y_2}$) assuming that the indicators of η ($\mathbf{y_1}$ $\mathbf{y_2}$) become causes of ξ; Mode (A,A) updated the estimates of η_0 by single regressions of $\mathbf{v_1}$ $\mathbf{v_2}$ $\mathbf{v_3}$ on η_0, in contrast with the paths specified between LV's in the structural model (Fig.1).
- because the scores of LV's are performed by means of alternating simple and/or multiple regressions, the parameter estimates are not performed by maximizing a global optimum.

The example treated so far requires the employment of all available information (MV) coherent with the measurement model specified for all LV's in a structural model; if $\mathbf{y_i}$ is a manifest variable of an endogenous LV, there is not reason to treat it as a variable that defines the score of its underlying latent variable, but on the other hand if $\mathbf{y_i}$ is a manifest variable of an exogenous LV (linked with an endogenous LV in the structural equation) it's reasonable to use it in a linear combination to perform the scores of a LV (Mode (A,A) updates the estimates of ξ_0 by single regressions of $\mathbf{y_1}$ $\mathbf{y_2}$ on ξ_0).

4. A measurement model for a single LV

Because the only information about an LV regards its role in the model and nothing is known about its unit of measure, mean, variance, it seems rational to conceive an LV as an unobservable construct strictly closed to the manifest variables (proxy of LV) which generate it (MV's as *causes*) and which better fits (in terms of regression and explained variance) the scores of a set of variables which describe its effects (MV's as *indicators*).

Going back to the example of human capital, we can define **h** as the linear aggregation of variables $\mathbf{x_i}$ (parents' schooling, occupation, wealth, years of schooling, years of occupation, income and wealth lagged, age, sex, geographic region, race, educational qualification, field of occupation) that best predicts $\mathbf{y_1}$ (job income) and $\mathbf{y_2}$ (financial income), using at best the prior information (from GFI) linking the MV's.

This specification of **h** is depicted in the Path diagram of Fig 2:

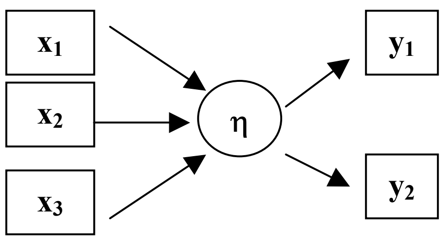

Fig.2: Measurement model for an LV with two sets of MV.

In presence of only LV η, the specification of a measurement model ([9]) for η requires the existence of two sets of MV's: **X** (n×p) contains causes, which generate its score, and **Y** (n×q) contains indicators of η:

$$\mathbf{Y} = \mathbf{1_n}\mu' + \mathbf{XAC} + \mathbf{E} \text{ with } \mathbf{A'X'XA=I}_r \;\; r(\mathbf{A})=r \;\; \text{Vec}(\mathbf{E}) \sim (\mathbf{0},\Sigma\otimes \mathbf{I_n}) \quad (2)$$

$$\mathbf{Y} = \mathbf{1_n}\mu' + \eta\,\mathbf{C} + \mathbf{E} \quad \text{with } r(\eta)=r \quad (3)$$

where **E** is the errors matrix, r (<min(p,q)) the rank of η(n×r), $\mathbf{1_n}$ (n×1) an unit vector, μ (q×1) the vector of **Y** means, Vec the usual operator which stacks columns. The model in (3), with the constraint that guarantees the uniqueness of **A** (p×r) in the Singular Value Decomposition, differently from the methodologies working in two

separate steps (PLS, RCD ([12])), forms a model for the simultaneous estimate of the weights **A** that define the LV η (η=**XA**) conceived as linear combinations of **X** (to avoid the non uniqueness of LV) and the regression weights **C** *(r×q)* between an LV and its indicators **Y** (structural parameter); finally the model (3) allows the association among $\mathbf{y_i}$, with the structured errors matrix Σ *(q×q)*.
With Σ previously estimated (by SURE Analysis [18], or by Covariance Structure Analysis [6]) the parameters estimates in model (2) are:

$$\mathbf{C} = \mathbf{A}' \mathbf{V}'\mathbf{Y} \qquad \mathbf{A} = \mathbf{T}^{-1} \mathbf{F} \tag{4}$$

where **F** is the matrix of eigenvectors corresponding to the first r eigenvalues of $\mathbf{V}'\mathbf{Y}\Sigma^{-1}\mathbf{Y}'\mathbf{V}$ and **T** is the upper diagonal matrix performed by Gram-Schmitd decomposition on **V**: $\mathbf{V} = \mathbf{V}^* \mathbf{T}$ (with $\mathbf{V}^*$ orthogonal).
Joreskog and Goldberger ([7]) called MIMIC a specified model with an LV in presence of multiple causes (**X**) and multiple indicators (**Y**), but their approach is more restrictive because it is based on stringent assumptions about the unitary rank of **A**, the structure of Σ (supposed to be diagonal, following the factor model), and the (normal) distribution of errors.
The methodology proposed has several advantages: the parameters are estimated consistently with the supposed relations (in the model) between LV's-causes, LV's-indicators, causes-indicators, avoiding treating the causes as indicators and/or vice versa, the structural parameters of measurement models are simultaneously estimated with the matrix of weights to define LV. The estimate of model (3) is extended to categorical and/or mixed manifest variables ([9], [10]).

5. Extension to two or more LV in a recursive structural model

In this section we extend the model (3) in presence of causal links between a set of endogenous η and exogenous ξ LV's .
The estimate of the structural model operates in two steps: in the first, we estimate the scores of LV's, in the second, the structural equations (Path Analysis) with the scores of estimated LV's.
The scores of LV's are performed by making rational use of prior information in the structural model: especially the paths between η and

ξ, and those linking the indicators of different latent variables.

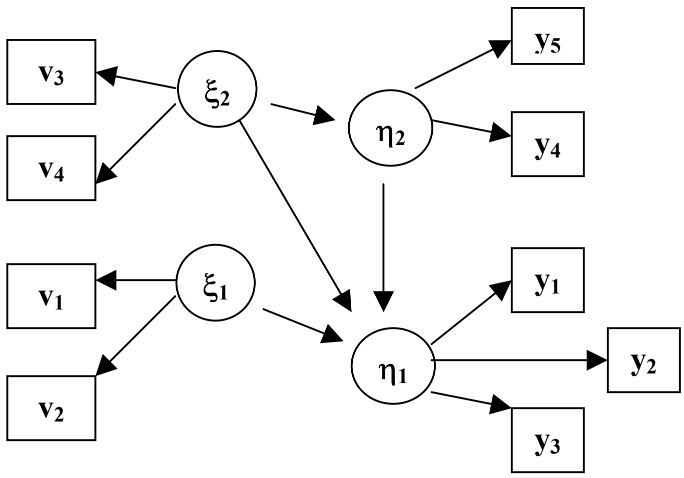

Fig.3: Path diagram of a recursive structural model

To extend the model proposed in Section 4 and its advantages, the first primary goal is to assign to each LV a set of causes and a set of indicators (following the terminology of MIMIC model).
The major drawback of the model proposed (es. in Fig.1) lies in the lack of variables **X** for each LV whereas the indicators are already specified in the model for all LV; nevertheless for the estimate of η_i the specification of paths between LV's suggests considering all the indicators of ξ_i (ξ_i are all exogenous LV of η_i) as the causes of η_i and estimate ξ_i as the linear combination of its indicators that best fits the indicators of η_i (where η_i are all endogenous LV in regards to ξ_i).
In the path model of Fig. 3 ξ_2 can be found as the linear combination of ($\mathbf{v}_3$, $\mathbf{v}_4$) that better fits ($\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$, $\mathbf{y}_4$, $\mathbf{y}_5$) because they are indicators of η_1 and η_2 endogenous LV in regards to ξ_2; ξ_1 can be viewed as the linear combination of ($\mathbf{v}_1$, $\mathbf{v}_2$) that better fits ($\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$); η_2 is the linear combination of ($\mathbf{v}_3$, $\mathbf{v}_4$) that better fits ($\mathbf{y}_4$, $\mathbf{y}_5$); η_1 is the linear combination of ($\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, $\mathbf{v}_4$, $\hat{\eta}_2$), with $\hat{\eta}_2$ the estimate of η_2 previously achieved, that better fits ($\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$).
For a general structural model:

$$\mathbf{H} = \mathbf{HB} + \Xi\Gamma + \mathbf{E} \quad \mathbf{Y} = \mathbf{H}\Lambda_y + \Delta \;\; ; \;\; \mathbf{V} = \Xi\Lambda_v + \mathbf{U} \qquad (5)$$

where $\mathbf{Y'}=(\mathbf{y'}_{(1)},\ldots,\mathbf{y'}_{(t)})$ (n_y,t), $\mathbf{V'}=(\mathbf{v'}_{(1)},\ldots,\mathbf{v'}_{(t)})$ (n_v,t) are the observed variables; $\Xi'=(\xi'_{(1)},\ldots,\xi'_{(t)})$ (n_ξ,t), $\mathbf{H'}=(\eta'_{(1)},\ldots,\eta'_{(t)})$ (n_η,t) are the latent variables; $\mathbf{E'}=(\varepsilon'_{(1)},\ldots,\varepsilon'_{(t)})$ (n_ε,t) are the errors in equations; $\Delta'=(\delta'_{(1)},\ldots,\delta'_{(t)})$ (n_δ,t), $\mathbf{U'}=(\mathbf{u'}_{(1)},\ldots,\mathbf{u'}_{(t)})$ (n_u,t), are the errors in variables; all the random variables have zero mean and finite variance, **B** is a low matrix with zero on the main diagonal**,** **(Y, V, H)** are identically distributed and (**Ξ,E,Δ,U**) are identically and independently distributed.
From the previous example, to achieve the estimate of LV's in a structural model with more η_j (with indicators in the matrix $\mathbf{Y}_j$) and ξ_i (with indicators in the matrix $\mathbf{V}_i$) we indicate the following rules to assign causes and indicators to each LV in (5):
1. if ξ_i is connected with only η_j (with MV's $\mathbf{Y}_j$) and η_j is not linked to others ξ_i $(i \neq j)$ the measurement models of η_j and ξ_i are indistinguishable: so $\hat{\xi}_i$ can be estimated as the first principal component of its MV's ($\mathbf{V}_i$) and the score of η_j is estimated as the combination of $\mathbf{V}_i$ that better fits $\mathbf{Y}_j$, by applying the equation (2).
2. if ξ_i is an exogenous LV (with MV's $\mathbf{V}_i$), directly connected with a set of η_i ..η_j ..η_k (j=1,..k) endogenous LV's (each with MV's $\mathbf{Y}_j$), the $\mathbf{V}_i$ are its causes and all variables in the blocks $\mathbf{Y}_j$ placed in the matrix $\mathbf{Y}= (\mathbf{Y}_1 \ldots \mathbf{Y}_k)$ its indicators: the estimate of ξ_i is performed from the model:

$$\mathbf{Y} = \mathbf{1}_n\mu' + \mathbf{V}_i\mathbf{A}\,\mathbf{C} + \mathbf{E} \;\; \Rightarrow \;\; \mathbf{Y} = \mathbf{1}_n\mu' + \xi_i\,\mathbf{C} + \mathbf{E} \quad \text{with} \;\; \xi_i'\xi_i = 1 \qquad (6)$$

3. if η_j is an endogenous LV, with MV's $\mathbf{Y}_j$, which is not endogenous in structural equations in regards to others endogenous LV η_k, (η_j is the first element in the **H** matrix of (4)), but η_j is directly connected with r exogenous LV's ξ_{ij} (i=1,..r, each with MV's $\mathbf{V}_{ij}$), the columns of $\mathbf{Y}_j$ are its indicators are and all variables in the blocks $\mathbf{V}_{ij}$ placed in the matrix $\mathbf{V}= (\mathbf{V}_{1j} \ldots \mathbf{V}_{rj})$ its causes:

$$\mathbf{Y}_j = \mathbf{1}_n\mu' + \mathbf{V}\mathbf{A}_j\,\mathbf{C} + \mathbf{E} \;\Rightarrow\; \mathbf{Y}_j = \mathbf{1}_n\mu' + \eta_j\,\mathbf{C} + \mathbf{E} \quad \text{with} \;\; \eta_j'\eta_j = 1 \qquad (7)$$

4. if η_j is an endogenous LV, with MV's $\mathbf{Y}_j$, in regards to a set of k LV's $\eta_1 \ldots \eta_k$ and directly connected with r exogenous LV's ξ_{ij} (i=1,..r, each

with MV $\mathbf{V_{ij}}$), the columns of $\mathbf{Y_j}$ are its indicators and the components of the matrix $\mathbf{V_j}$*=($\mathbf{V_{1j}} \ldots \mathbf{V_{rj}}$, **Z**), where **Z**=($\hat{\eta}_1 \ldots \hat{\eta}_k$) have been previously estimated, its causes:

$$\mathbf{Y_j = 1_n\mu' + V_j{}^*A\, C + E} \Rightarrow \mathbf{Y_j = 1_n\mu' + \eta_j\, C + E} \quad \text{with } \eta_j{}'\eta_j = 1 \qquad (8)$$

5. if η_j (with MV's $\mathbf{Y_j}$) is only endogenous in regards to a set of k LV's $\eta_1 \ldots \eta_k$ but not in regards to any ξ_i, the columns of $\mathbf{Y_j}$ are its indicators and the previously estimated components of a matrix **Z**=($\hat{\eta}_1 \ldots \hat{\eta}_k$) its causes:

$$\mathbf{Y_j = 1_n\mu' + ZA\, C + E} \quad \Rightarrow \quad \mathbf{Y_j = 1_n\mu' + \eta_j\, C + E} \qquad (9)$$

The previous models define the rules for the estimate of an LV as a linear combination of manifest variables that best fits (predicts) a set of indicators by using all available information between these two sets of MV's, following the paths between endogenous and exogenous LV, and allows the simultaneous estimate of scores η_j and regression parameter (**C**) between each endogenous LV and its indicators.
Following these rules every LV characterized by submatrices $l_{y(\beta,\bullet)}, l_{x(\delta,\bullet)}$ with coefficients all different from zero in the factor models of (5), is estimated ($\tilde{\eta}_\beta$ and $\tilde{\xi}_\delta$).

6. Constraints on parameters

Then in order to take into account the restrictions of a structural model:

$$\mathrm{cov}(\eta_\beta, \eta_\pi) = 0; \mathrm{cov}(\xi_\delta, \xi_\gamma) = 0; \mathrm{cov}(\delta_{\beta\mathbf{k}}, \delta_{\beta\vartheta}); \mathrm{cov}(\mathbf{u}_{\delta\lambda}, \mathbf{u}_{\delta\mu}) = 0;$$
$$b_{(\beta,\mu)} = 0; \gamma_{(\delta,\varphi)} = 0; \mathbf{l}_{\mathbf{y}(\beta,\alpha)} = 0; \mathbf{l}_{\mathbf{v}(\delta,\nu)} = 0; \qquad (10)$$

where the first row lies with the hypothesized incorrelations between the parameters in (5) and the second row constraints submatrices in the structural model and two measurement models to be null, respectively.
By means of the Restricted Regression Component Decomposition (RRCD [4], [5]) of LV $\tilde{\eta}_\beta$ and $\tilde{\xi}_\delta$ by means of an iterative process:

$$\eta_\beta^+ = Q_{*\eta_\mu \cup *y_\alpha}\,\eta_\beta^o\ (\eta_\beta^0 = Q_{\tilde{\eta}_\pi}\tilde{\eta}_\beta\,(\beta \neq \pi);{*}\,\eta_\mu = Q_{H^o_{(\beta,\mu)}}\,\eta_\mu^o;{*}\,y_\alpha = Q^o_{H(\beta)}\,y_\alpha)$$
$$\xi_\delta^+ = Q_{*\eta_\varphi \cup *x_\nu}\,\xi_\delta^o\ (\xi_\delta^0 = Q_{\tilde{\xi}_\gamma}\tilde{\xi}_\delta\,(\gamma \neq \delta);{*}\,\eta_\varphi = Q_{H^o_{(\varphi)} \cup \Xi^o_{(\delta)}}\,\eta_\varphi^o;{*}\,x_\nu = Q_{\Xi^o_{(\delta)}}\,x_\nu)$$
$$\tilde{\varepsilon}_j = Q_{\tilde{H}(j) \cup \tilde{\Xi}}\,\tilde{\eta}_j\ ((j)=1,\dots,j\text{-}1);$$
$$\delta^0_{\beta k} = Q_{\left(H^0 \cup \Xi^0 \cup y_{\beta_\vartheta}\right)}\,y_{\beta k};$$
$$u^0_{\delta\lambda} = Q_{\left(X \cup \Xi^0 \cup x_{\delta_\mu}\right)}\,x_{\delta\lambda} \qquad (11)$$

where $H^\circ_{(\mu,\beta)}$ are the H° without $\eta^\circ_\mu, \eta^\circ_\beta$, $Q_{\tilde{\eta}_\pi}$ is the complement orthogonal to the orthogonal projector on the space generated by $\tilde{\eta}_\pi$ and the other symbols are defined in a similar way. So we have the following RRCD of $y_{\beta k}, x_{\delta\lambda}, \eta^o_\beta$:

$$Q_{\Xi^0 \cup y^0_{\beta_\vartheta}/H^0}\,y_{\beta_k} = P_{H^0}\,y_{\beta_k} + Q_{H^0 \cup \Xi^0 \cup y_{\beta_\vartheta}}\,y_{\beta_k};$$
$$Q_{Y \cup x_{\delta_\mu}/\Xi^0}\,x_{\delta_\lambda} = P_{\Xi^0}\,x_{\delta_\lambda} + Q_{Y \cup x_{\delta_\mu} \cup \Xi^0}\,x_{\delta_\lambda} \qquad (12)$$

$$\eta^o_\beta = P_{H^0_{(\beta)}}\,\eta^o_\beta + P_{\Xi^0/H^0_{(\beta)}}\,\eta^o_\beta + Q_{H^0_{(\beta)} \cup \Xi^0}\,\eta^o_\beta$$

7. Application

The model (3) is applied for the estimate of the scores of human capital (**h**) conceived as a bidimensional LV (schooling and professional experience on the job): Tab.1 specifies for each dimension a set of different causes and the same set of indicators (financial income and job income), following basically the GFI and the recursive model specified by Dagum ([1], [2]). The analysis is based on 4103 American Families (Federal Reserve Board Survey, 1983). The parameter estimates follow in Tab.1 and the scores of each LV are depicted in Fig.4.
The redundancy index ([9]), which can be taken to assess the goodness of model fit and like optimum criteria, results 0.7334.

Variables	Description*	Coefficients	
		Schooling	Esperience
Intercept		-162,427	-10,844
X1	Age H	-,005	
X2	Sex H (1=Female)	3,451	
X5	Marital status H	-3,243	
Y1	Years of schooling H	1,040	0,616
Y2	Years of schooling S	1,057	
Y4	Years of job experience H	0,053	0,210
Y6	Years of job experience S		-,0742
Y8	Job status H		-2,231
Y9	Occupation H		-4,824
Y10	Job Sector H		-4,477
Y11	Job status S		0,877
Y14	Total wealth F	3,021	

*data are for Families (F), Head (H), Spouse (S) For better reference see [9].

Tab. 1: Variables, description , and parameter estimation

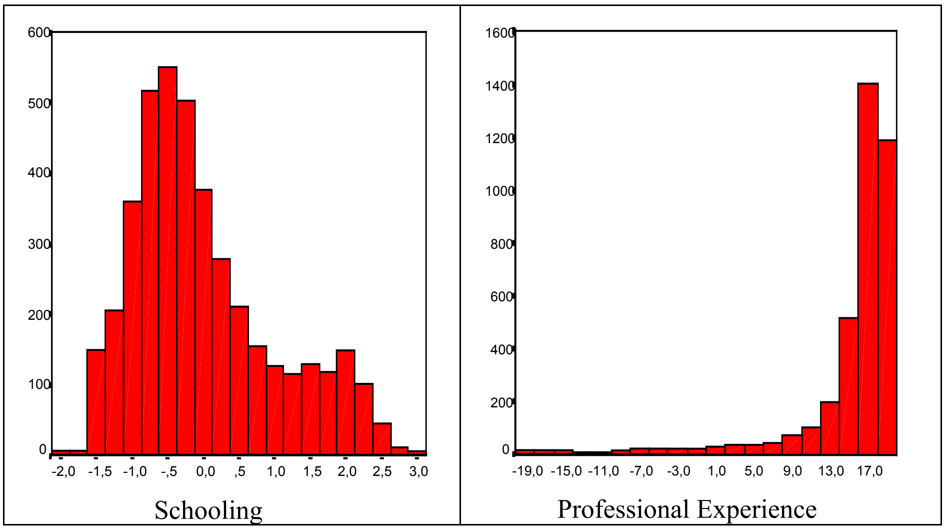

Fig.4: Distribution of scores of estimated LV

7. Conclusion

The measurement model for a single LV and two sets of MV's with a logically different link with a LV is proposed. The **X** variables for an endogenous LV can be taken from the paths between LV's, avoiding the arbitrary introduction of new **X** variables and the inversion of the causal roles between MV's and LV's ([9]) or between exogenous and endogenous LV's ([15]), as PLS methodology does.

A criticism of PLS ([3]) is that there seems to be no well-defined modeling problem for which it provides the optimal solution, other than the arbitrary criteria to optimize: the model and the rules proposed in this paper can be useful for PLS users to better understand the performed scores of LV's, because they are achieved by the same optimum criteria.

The model (3) is a true causal model (like the factor model) because the linear combination that defines the scores of LV are performed in a statistical model and the error matrix is interpretable as stochastic errors or as errors in equations.

Finally, to take into account the relations specified by the structural model between MV's in the same block of LV in the factor models, or between LV's in the structural model, it's possibile to apply a RCDR analysis starting from the previously estimated scores of LV's: the estimated parameter matrices satisfy all the properties of the LV's in the structural model, contrary to those of the PLS ([15]), and respect the constraints on the parameters, summarized in (10).

BIBLIOGRAPHY

[1] Dagum C. (1994): Human Capital, Income and Wealth Distribution Models and Their Applications to the USA, *Proceedings of the 154th Meeting of the American Statistical Association*.

[2] Dagum C. Vittadini G. (1996): Human Capital Measurement and Distribution, *Proceedings of the 156th Meeting of the American Statistical Association, Business and Economic Statistics Section*.

[3] Garthwaite P.H. (1994): An interpretation of Partial Least Squares, *Journal of American Statistical Association*, 89, 122-127.

[4] Haagen, K., Vittadini, G. (1998): Regression Component Decomposition Restricted. Un'alternativa al Lisrel Model, *Metron*, 56, 1-12.

[5] Haagen K., Vittadini G. (1991): Regression Component Decomposition in Structural Analysis, *Communications in Statistics*, 20, 1153-1161.

[6] Joreskog K. (1978): Structural Analysis of Covariance and Correlation matrices, *Psychometrika*, 43, 443-477.

[7] Joreskog K., Goldberger S. (1975): Estimation of a model with Multiple Indicators and Multiple Causes of a single Latent Variable, *Journal of American Statistical Association*, 70, 631-639.
[8] Joreskog K., Wold H. (1982): *Sistem under indirect observation*, North Holland, Amsterdam.
[9] Lovaglio P.G. (2001): La stima di Variabili Latenti da variabili osservate miste, *Statistica*, in press.
[10] Lovaglio P.G. (2001): Un algoritmo di regressione multipla con dati misti. *Quaderni di Matematica e Statistica Applicata alle Scienze Socio-Economiche*, Trento, in press.
[11] Schneeweiss H. (1991): Models with latent variables: LISREL versus PLS, *Statistica Neederlandica* 45, 1454-157
[12] Schonemann P., Steiger J. (1976): Regression Component Analysis, *British Journal of Mathematical and Statistical Psychology,* 29, 175-189.
[13] Tenenhaus M. (1995) *La Régression PLS: Théorie et Pratique*. Editions Technip, Paris.
[14] Vittadini G. (1999): Analysis of qualitative variables in Structural Models with unique solutions. In: M. Vichi O. Opitz (eds.), *Classifications and data analysis-Theory and Applications*, Springer and Verlag, 203-210.
[15] Vittadini G. (1992): Un confronto tra un modello di analisi causale con variabili latenti e metodi alternativi, *Statistica*, LII, 3, 379-396.
[16] Vittadini G. (1989): Indeterminacy problems in the Lisrel Model, *Multivariate Behavioral Research*, 24, 397-414.
[17] Wold H. (1982): Soft modeling: the basic design and some extension. In: K. Joreskog Wold H. (eds.), *Sistem under indirect observation*, North Holland, Amsterdam.
[18] Zellner A. (1962): An efficient method of estimating Seemingly Unrelated Regressions and test for aggregation bias, *Journal of Americal Statistical Association*, 348-367.