METRON

RIVISTA INTERNAZIONALE DI STATISTICA — REVUE INTERNATIONALE DE STATISTIQUE
INTERNATIONAL JOURNAL OF STATISTICS — INTERNATIONALE STATISTISCHE ZEITSCHRIFT
REVISTA INTERNACIONAL DE ESTADISTICA

FOUNDED IN 1920 BY
Corrado Gini

Vol. LVI n. 1-2 1998

K. HAAGEN - G. VITTADINI

Restricted regression component decomposition

DIPARTIMENTO DI STATISTICA, PROBABILITA' E STATISTICHE APPLICATE
UNIVERSITA' DEGLI STUDI DI ROMA «LA SAPIENZA»

KLAUS HAAGEN (*) - GIORGIO VITTADINI (**)

Restricted regression component decomposition

1. Introduction

The identification of the parameters in the models with latent variables, where a causal relation is assumed between the observed and the latent variables (Jöreskog 1970, 1971, 1973a, 1973b, 1974, 1977, 1978, 1979a, 1979b, 1981a, 1981b, 1982), is an unsolved problem (Jöreskog 1981b, Everitt 1984). However, even if the parameters are completely identified, another identification problem in the model remains: the indeterminacy of the latent variables (Vittadini 1988, 1989).

In order to resolve the identification and the indeterminacy problems, alternative methods have been proposed, such as the Partial Least Squares (PLS) (Wold 1980, Apel and Wold 1982, Noonan and Wold 1982), and the extension of the Regression Component Decomposition (RCD) (Schönemann and Steiger, 1976, Haagen and Vittadini, 1991) to the structural models.

This paper proposes a structural model that obtains the latent variables in terms of linear transformation from their observed indicators in a flexible way with respect to *a priori* restriction and information.

(*) Università degli Studi di Trento, Via Inama, 5 - 38100, Italia.
(**) Università degli Studi di Milano, Viale Sarca, 202 - 20136, Italia.

2. The structural model and previous proposals

The structural model is composed of one structural and two measurement equations:

$$\eta_{(t)} = B'\eta_{(t)} + \Gamma'\xi_{(t)} + \varepsilon_{(t)} \quad (1a)$$

$$y_{(t)} = \Lambda_Y'\eta_{(t)} + \delta_{(t)} \quad (1b)$$

$$x_{(t)} = \Lambda_X'\xi_{(t)} + u_{(t)} \qquad t = 1, \ldots T \quad (1c)$$

It is assumed that all the random variables have zero mean and finite variance, and that B' is a triangular matrix with zero on the main diagonal. $y_{(t)}, x_{(t)}, \eta_{(t)}, \xi_{(t)}$ $(t = 1, \ldots, T)$ are vectors respectively with n_y, n_x, n_η, n_ξ components $(n_y > n_\eta;\ n_x > n_\xi,\ T > n_X + n_y)$ identically distributed and:

$\xi_{(t)}, \varepsilon_{(t)}, \delta_{(t)}, u_{(t)}$ $(t = 1, \ldots, T)$ are identically and independently distributed;

(U, Δ, E) and Ξ independent;

(U, Δ) and E independent;

U and Δ independent.

where: $\Xi' = (\xi_{(1)}, \ldots, \xi_{(T)})$, $E' = (\varepsilon_{(1)}, \ldots, \varepsilon_{(T)})$, $\Delta' = (\delta_{(1)}, \ldots, \delta_{(T)})$, $U' = (u_{(1)}, \ldots, u_{(T)})$.

It is useful to write (1) in the form:

$$z = A_{z\psi}\psi + \theta \quad (2)$$

$$\text{with } z = \begin{bmatrix} y \\ x \end{bmatrix} \psi = \begin{bmatrix} \xi \\ \varepsilon \end{bmatrix}, \theta = \begin{bmatrix} \delta \\ u \end{bmatrix}, A_{z\psi} = \begin{bmatrix} A_{y\xi} & A_{y\varepsilon} \\ A_{x\xi} & A_{x\varepsilon} \end{bmatrix} = \begin{bmatrix} \Lambda'_y(I-B')^{-1}\Gamma', & \Lambda'_y(I-B')^{-1} \\ \Lambda'_X, & 0 \end{bmatrix}$$

If $H' = (\eta_{(1)}, \ldots, \eta_{(T)})$; $X' = (x_{(1)}, \ldots, x_{(T)})$; $Y' = (y_{(1)}, \ldots, y_{(T)})$ are the observations of the variables, then we have:

$$H = HB + \Xi\Gamma + E \quad (3a)$$

$$Y = H\Lambda_y + \Delta \quad (3b)$$

$$X = \Xi\Lambda_X + U \quad (3c)$$

For the variance-covariance matrix we get:

$$\Sigma_{ZZ} = A_{Z\psi}\Sigma_{\psi\psi}A'_{Z\psi} + \Sigma_{\Theta\Theta} \quad (4)$$

$$\text{where: } \Sigma_{\psi\psi} = \begin{bmatrix} \Sigma_{\Xi\Xi} & 0 \\ 0 & \Sigma_{EE} \end{bmatrix} \ (\Sigma_{\Xi E} = 0).$$

The solutions are usually obtained from (4) as latent causes of the variations and covariations of the observed variables Z, through estimation methods such as maximum likelihood, or the least squares (unweighted or weighted), and successively through methods of numerical calculation.

However, these solutions in general are not unique. In effect, in the case of the Lisrel model "no general and practically useful necessary and sufficient conditions for identification are available" (Everitt 1984).

In practice, if we do not consider a few very restrictive cases in which conditions for identifiability are studied analytically [e.g. where the endogenous variables are measured without error ($H = Y$) (Geraci 1976)], it is suggested that "the identification problem be studied on a case by case basis by examining the equations" Jöreskog (1981b), choosing the restrictions, not only in number but also in position, in order to obtain unique solutions. This is also true in the case of local identifiability of the parameters (Fisher 1966, Rothenberg 1971, Geraci 1976, Bekker and Pollock 1985, Shapiro 1985, Bekker 1989, Wegge 1965, Wegge 1991, Wegge 1996).

The problem becomes more complex when one uses, as in experimental sciences (Wold 1982, Bentler 1982), the model with *a priori* information given by empirical problems. In fact, in these disciplines, the restrictions are a set of hypotheses inherent to the relationships between the observed variables, the latent variables and the errors. Therefore, every *a priori* hypothesis relative to a particular specification of the model generated by particular real situations can be very different, even in opposition in certain cases, to the imposed necessary "case by case" (Jöreskog 1981b) mathematical conditions used to identify the model.

When the parameter identifiability is given, the model does not supply unique scores for the latent variables, because it is not possible to obtain from the space generated by the columns of observed variables a unique space of latent variables and errors of greater dimensions (Guttmann 1955). It is also shown that the indeterminacy considered more or less harshly in the literature (Mulaik 1972, Schönemann and Wang 1972, McDonald 1978, Schönemann and Steiger 1978, Haagen 1986, 1987, 1991, Williams 1978, Bentler 1976, 1982; Schönemann and Haagen 1987) does not only regard the scores of the latent variables but also implies the existence of logically inconsistent properties for the whole model (Vittadini 1988, 1989).

3. The alternative proposals

To avoid the problems presented thus far, other methods for finding solutions for structural models have been proposed. Such methods obtain latent variables, such as linear combinations of the observed variables and residual errors. Such variables cannot be precisely understood as latent variables if one accepts the definition of Bentler (1982) who states that "a necessary and sufficient condition for a linear structural equation system to be a latent variable model is that the dimensionality of space spanned by the independent variables is greater than the dimensionality of the space spanned by the manifest variables". The same author, however, summarizing contributions of a set of authors (Dempster 1971, Schönemann and Steiger 1976, Wold 1980), recognizes that the methodologies based on linear combinations can well be used to evaluate the latent variables because "there is no empirical way of distinguishing between the models", and therefore "it is possible to use the observed variables as a vehicle for estimating parameters of the latent variables model" (Bentler 1982). In order to be verified, it is necessary that the solutions based on linear combinations have the same properties as the latent variables required by the structural models. This does not occur in the case of the PLS (Wold 1982, Noonan and Wold 1982) when the weights of the linear combinations are obtained through an iterative procedure based on multiple regression of the observed variables on the latent variables or on simple regression of the latent variables on the observed variables (Djikstra (1983) described the stochastic properties of these PLS estimates).

In fact, there is a lack of objective criteria (Noonan and Wold 1982) in the choice of weights for the latent variables; moreover, solutions assumed by the PLS do not respect all the uncorrelations among the latent variables and errors, and thus it becomes impossible to define the matrix Λ_y, Λ_x, B, Γ as regression matrices (Vittadini 1992), as required in the structural models.

With the proposal of Haagen-Vittadini (1991), the extension of the RCD valid for the factorial model (Schönemann and Steiger 1976) through a decomposition of the space of the variables Z in the model expressed in reduced form (2) obtains unique solutions for the latent variables and the parameter matrices of the model.

All the uncorrelations are respected, the parameter matrices are regression matrices perfectly identified, and the scores of the latent variables are unique.

Two problems remain unsolved. First, the proposal of Haagen and Vittadini ignores *a priori* restriction and information, while it seems that such models are applied in situations where we have *a priori* knowledge.

Another problem is connected to the fact that the solutions of Haagen and Vittadini, like Jöreskog's solutions, are obtained by starting from a reduced model; in the first case, starting from the model expressed in terms of observed variables scores (2); in the second, from the model expressed in terms of the variance-covariance matrix (4). It can be shown (Schneeweiß 1991) that the values obtained for reduced model solutions are different from those obtained by directly using the measurement models (1b), (1c). It is reasonable to hypothesize that because of the nature of latent variables (Kmenta 1991), particularly because we must treat theoretical constructions as we do structural models (Schneeweiß 1991), it is preferable to obtain them from their direct observed indicators.

4. The new proposal: restricted regression component decomposition

The method to determine the parameters and the latent scores presented in this paper is not an estimation procedure but a method to define latent variables and to derive their relations to each other from covariance structure and draft variables.

As in PLS, latent variables are defined as linear transformation of their direct indicators which are most appropriate to the particular goal of the analyses, and most flexible on the basis of different restrictions and information. Following this, three proposals are put forward:

a) a model without *a priori* restrictions and information on the variables;
b) a model with *a priori* information on the variables and on the errors;
c) a model with *a priori* parameter restrictions and information.

4.1. *Model without restrictions*

Starting from (3b) (3c), we can define, as in the case of Wold's model, the subsets of observed variables characterized by submatrices of parameters $l_{x(\dots,\dots)}$ e $l_{y(\dots,\dots)}$, which all have elements different from zero:

$$Y_\alpha = \tilde{\eta}_\alpha l'_{y(.,\alpha)} + \Delta_\alpha \quad (\alpha = 1, \dots, n_\eta) \qquad X_\gamma = \tilde{\xi}_\gamma l'_{x(.,\gamma)} + U_\gamma \quad (\gamma = 1, \dots, n_\xi) \quad (5)$$

with $Y_\alpha(T, n_\alpha)$, $\tilde{\eta}_\alpha(T, 1)$, $l_{y(.,\alpha)}(n_\alpha, 1)$, $\Delta_\alpha(T, n_\alpha)$, $X_\gamma(T, n_\gamma)$, $\tilde{\xi}_\gamma(T, 1)$, $l_{x(.,\gamma)}(n_\gamma, 1)$, $U_\gamma(T, n_\gamma)$

and one obtains η_α and ξ_γ by means of linear transformations:

$$\tilde{\eta}_\alpha = Y_\alpha k_\alpha \quad (\alpha = 1, \dots, n_\eta); \qquad \tilde{\xi}_\gamma = X_\gamma w_\gamma \quad (\gamma = 1, \dots, n_\xi) \quad (6)$$

when $k_\alpha(n_\alpha\ 1)$, $w_\gamma(n_\gamma\ 1)$ are parameter vectors.

In some cases, subsets of observed variables relative to different latent variables, $\tilde{\eta}_\alpha$, $\tilde{\eta}_\beta$ that have one or more indicators in common, are obtained in the following way:

$$\tilde{\eta}_\alpha = Y_\alpha^{**} k_\alpha; \qquad \tilde{\eta}_\beta = Y_\beta^{*} k_\beta \quad (7)$$

with $y^{**}_{\beta_j} = Q_{\tilde{\eta}_\beta} y_{\beta_j}$ in Y^{**}_α, $y^{*}_{\beta_j} = Q_{\tilde{\eta}_\alpha} y_{\beta_j}$ in Y^{*}_β, the indicators in common between $\tilde{\eta}_\alpha$ and $\tilde{\eta}_\beta$, where $P_{\tilde{\eta}_\alpha}$ and $P_{\tilde{\eta}_\beta}$ are the orthogonal projectors respectively on the spaces generated by the columns of $\tilde{\eta}_\alpha$ and $\tilde{\eta}_\beta$ (Takeuchi et al. 1982), $Q_{\tilde{\eta}_\alpha} = I - P_{\tilde{\eta}_\alpha}$ and $Q_{\tilde{\eta}_\beta} = I - P_{\tilde{\eta}_\beta}$ the orthogonal complements to the spaces generated by the columns of $\tilde{\eta}_\alpha$ and $\tilde{\eta}_\beta$. In the same way, if subsets of observed indicators relative to different latent variables $\tilde{\xi}_\gamma$ and $\tilde{\xi}_\delta$ are not disjoint, we obtain:

$$\tilde{\xi}_\gamma = X_\gamma^{**} w_\gamma; \qquad \tilde{\xi}_\delta = X_\delta^{*} w_\delta \quad (8)$$

with $x^{**}_{\gamma_j} = Q_{\tilde{\xi}_\delta} x_{\gamma_j}$ in X^{**}_γ; $x^{*}_{\gamma_j} = Q_{\tilde{\xi}_\gamma} x_{\gamma_j}$ in X^{*}_δ

At this point, applying the usual RCD, one can obtain from (3a) the matrices B, Γ and errors E. It is important, however, to define how the parameter matrices coherent with the hypotheses of the structural model are obtained by this method. Starting with (1a) and (3a) we get:

$$\tilde{\eta}_j = P_{\tilde{H}_{(j)}} \tilde{\eta}_j + P_{\tilde{\Xi}/\tilde{H}_{(j)}} \tilde{\eta}_j + Q_{\tilde{H}_{(j)} \cup \tilde{\Xi}} \tilde{\eta}_j; \quad V_{\tilde{\eta}_j} = V_{\tilde{H}_{(j)}} \oplus V_{\tilde{\Xi}/\tilde{H}_{(j)}} \oplus V_{\tilde{\eta}_j/\tilde{H}_{(j)} \cup \tilde{\Xi}} \quad (9)$$

where $\tilde{H}_{(j)}$ is the matrix $\tilde{H}_{(j)} = [\tilde{\eta}_1, \tilde{\eta}_2, \dots, \tilde{\eta}_{j-1}]$, $P_{\tilde{\Xi}/\tilde{H}_{(j)}}$ the orthogonal projection on the space generated by the columns of $\tilde{\Xi}$ orthogonal to the space generated by the columns of $\tilde{H}_{(j)}$, $Q_{\tilde{H}_{(j)} \cup \tilde{\Xi}}$ the orthogonal complement to the space generated by the columns of $\tilde{H}_{(j)} \cup \tilde{\Xi}$; and where $V_{\tilde{\eta}_j}$, $V_{\tilde{H}_{(j)}}$ are respectively the spaces generated by the columns of $\tilde{\eta}_j$ and by the columns of $\tilde{H}_{(j)}$, $V_{\tilde{\Xi}/\tilde{H}_{(j)}}$ the space generated by the columns of $\tilde{\Xi}$ orthogonal to the space generated by the columns of $\tilde{H}_{(j)}$, $V_{\tilde{\eta}_j/\tilde{H}_{(j)} \cup \tilde{\Xi}}$ the space generated by the $\tilde{\eta}_j$ orthogonal to the space generated by the columns of $\tilde{H}_{(j)} \cup \tilde{\Xi}$ (when $\oplus$ indicates the orthogonal sum of vector spaces).

From (9) one deduces that:

$$\tilde{\varepsilon}_j = Q_{\tilde{H}_{(j)} \cup \tilde{\Xi}} \tilde{\eta}_j \qquad (j > (j)) \quad (10)$$

For the partial regression coefficient $b_{j\varphi}$ of B:

$$b_{j\varphi} = \tilde{\eta}'_j Q_{\tilde{H}_{((j),(\varphi))}} \eta_\varphi / \tilde{\eta}'_\varphi Q_{\tilde{H}_{(\varphi)}} \tilde{\eta}_\varphi \qquad [\varphi = 1, \dots, n_{\eta-1};\ \varphi < j;\ ((j), (\varphi)) \neq (j, \varphi)] \quad (11)$$

and for the partial regression coefficient $\gamma_{j\eta}$ of Γ:

$$\gamma_{j\mu} = \tilde{\eta}'_j Q_{\tilde{H}_{(j)}} Q_{\tilde{\Xi}_{(\mu)}/\tilde{H}_j} Q_{\tilde{H}_{(j)}} \tilde{\xi}_\mu / \tilde{\xi}'_\mu Q_{\tilde{H}_{(j)}} Q_{\tilde{\Xi}_{(\mu)}/\tilde{H}_{(j)}} Q_{\tilde{H}_{(j)}} \tilde{\xi}_\mu =$$

$$= \tilde{\eta}'_j Q_{\tilde{H}_{(j)}} \left[I - P_{\tilde{\Xi}_{(\mu)}/\tilde{H}_{(j)}} - P_{\tilde{H}_{(j)}} \right] Q_{\tilde{H}_{(j)}} \tilde{\xi}_\mu / \tilde{\xi}'_\mu Q_{\tilde{H}_{(j)}} \left[I - P_{\tilde{\Xi}_{(\mu)}/\tilde{H}_{(j)}} - P_{\tilde{H}_{(j)}} \right] Q_{\tilde{H}_{(j)}} \tilde{\xi}_\mu =$$

$$= \tilde{\eta}'_j Q_{\tilde{H}_{(j)} \cup \tilde{\Xi}_{(\mu)}} \tilde{\xi}_\mu / \tilde{\xi}'_\mu Q_{\tilde{H}_{(j)} \cup \tilde{\Xi}_{(\mu)}} \tilde{\xi}_\mu \quad (12)$$

where $\tilde{H}_{(j,\varphi)}$ are the $\tilde{H}$ different from $\tilde{\eta}_j$ and $\tilde{\eta}_\varphi$ and $\tilde{\Xi}_{(\mu)}$ are the $\tilde{\Xi}$ different from ξ_μ. The partial regression coefficient $b_{j\varphi}$ component of b_j indicates the links between the $\tilde{\eta}_j$ and the $\tilde{\eta}_\varphi$ free from every influence of $\tilde{\Xi}$ and $\tilde{H}_{((j),(\varphi))}$ while the partial regression coefficient $\gamma_{j\mu}$ (identical to the analogous partial regression coefficient $f_{j\mu}$ of $F_j = [\tilde{H}_{(j)}, \tilde{\Xi}]$ on $\tilde{\eta}_j$), indicates the link between the η_j and the ξ_μ free from the influence of all the other independent latent variables $\tilde{H}_{(j)}$ and $\tilde{\Xi}_{(\mu)}$ with respect to the already stated hypotheses of the structural model (Jöreskog 1981b, Bentler and Weeks 1980, Bentler 1982, 1983). But the definition (10) determines the uncorrelation of every error $\tilde{\varepsilon}_j$ with every other $\tilde{\varepsilon}_k$ ($j \neq k$). In fact:

$$\text{cov}(\varepsilon_j, \varepsilon_k) = [Q_{\tilde{H}_{(j)} \cup \tilde{\Xi}}\tilde{\eta}_j, Q_{\tilde{H}_{(k)} \cup \tilde{\Xi}}\tilde{\eta}_k] = 0$$

$$(j > (j),\ k > (k),\ k \neq j) \tag{13}$$

At this point, without the uncorrelation hypothesis of the model, the RCD of the measurement models would be:

$$V_Y = V_{\tilde{H}} \oplus V_{Y/\tilde{H}};\ Y = P_{\tilde{H}}Y + Q_{\tilde{H}}Y;\ V_X = V_{\tilde{\Xi}} \oplus V_{X/\tilde{H}};\ X = P_{\tilde{\Xi}}X + Q_{\tilde{\Xi}}X \tag{14}$$

with
$$\Delta = Q_{\tilde{H}}Y; \qquad U = Q_{\tilde{\Xi}}X \tag{15}$$

Nevertheless, in order to satisfy (3) one needs to define in the first measurement model the errors Δ in the subspace of Y orthogonal to the space generated by the columns of $\tilde{H}$ and $\tilde{\Xi}$. Analogously, in the second model of measurement, the errors U are defined in the subspace of X orthogonal to the space generated by the columns of Y and $\tilde{\Xi}$:

$$\tilde{\Delta} = Q_{\tilde{H} \cup \tilde{\Xi}}Y; \qquad \tilde{U} = Q_{Y \cup \tilde{\Xi}}X \tag{16}$$

The new definitions of errors (16) determine some modification of the space on which the RCD (14) is carried out to respect the properties of the model. We have:

$$Q_{\tilde{\Xi}/\tilde{H}}Y = [I - P_{\tilde{\Xi}/\tilde{H}}]\,P_{\tilde{H}}Y + Q_{\tilde{\Xi} \cup \tilde{H}}Y = P_{\tilde{H}}Y + Q_{\tilde{\Xi} \cup \tilde{H}}Y; \tag{17a}$$

$$Q_{Y/\tilde{\Xi}}X = [I - P_{Y/\tilde{\Xi}}]\,P_{\tilde{\Xi}}X + Q_{Y \cup \tilde{\Xi}}X = P_{\tilde{\Xi}}X + Q_{Y \cup \tilde{\Xi}}X \tag{17b}$$

where we see that no changes occur on the parts of RCD relative to the variables $\tilde{H}$ and $\tilde{\Xi}$ in the first and second measurement models.

Therefore, we obtain the following measurement models:

$$V_{Q_{\tilde{\Xi}/\tilde{H}}Y} = V_{\tilde{H}} \oplus V_{Y/\tilde{H}\cup\tilde{\Xi}}; \qquad Q_{\tilde{\Xi}/\tilde{H}}Y = P_{\tilde{H}}Y + Q_{\tilde{\Xi}\cup\tilde{H}}Y \tag{18a}$$

$$V_{Q_{Y/\tilde{\Xi}}X} = V_{\tilde{\Xi}} \oplus V_{X/Y\cup\tilde{\Xi}}; \qquad Q_{Y/\tilde{\Xi}}X = P_{\tilde{\Xi}}X + Q_{Y\cup\tilde{\Xi}}X \tag{18b}$$

The Restricted Regression Component Decomposition (RRCD) of the first measurement model is equivalent to the RCD in the space $V_{Q_{\tilde{\Xi}/\tilde{H}}Y}$ obtained by reducing the space V_Y (14) to the projection on V_Y of the orthogonal complement of the space generated by the columns of $\tilde{\Xi}$ orthogonal to the space generated by the columns of $\tilde{H}$; the RRCD of the second measurement model is equivalent to the RCD in the space $V_{Q_{Y/\tilde{\Xi}}X}$, obtained by reducing the space V_x (14) to the projection on V_X of the orthogonal complement of the space generated by the columns of Y orthogonal to the space generated by the columns of $\tilde{\Xi}$.

From (18), one deduces that:

$$\Lambda_Y = \Sigma_{\tilde{H}}^{-1}\tilde{H}'Y; \qquad \Lambda_X = \Sigma_{\tilde{\Xi}}^{-1}\tilde{\Xi}'X \tag{19}$$

Therefore, with the RRCD, given the observed variables Y and X, the parameter matrices Λ_Y, Λ_X, B, Γ are identified and the latent variables and errors $\tilde{H}$, $\tilde{\Xi}$, $\tilde{E}$, $\tilde{U}$, $\tilde{\Delta}$ are unique.

Moreover, the latent variables and errors of the model respect all the properties of uncorrelation.

In fact, one has from (6), (10), (16):

$$\text{cov}(\tilde{\Xi}, \tilde{\varepsilon}_j) = \text{cov}(\tilde{\Xi}, Q_{\tilde{H}_{(j)}\cup\tilde{\Xi}}\tilde{\eta}_j) = 0;$$

$$\text{cov}(\tilde{\Xi}, \tilde{\Delta}) = \text{cov}(\tilde{\Xi}, Q_{\tilde{H}\cup\tilde{\Xi}}Y) = 0;$$

$$\text{cov}(\tilde{\Xi}, \tilde{U}) = \text{cov}(\tilde{\Xi}, Q_{Y\cup\tilde{\Xi}}X) = 0;$$

$$\text{cov}(\tilde{\varepsilon}_j, \tilde{\Delta}) = \text{cov}(Q_{\tilde{H}_{(j)}\cup\tilde{\Xi}}\tilde{\eta}_j, Q_{\tilde{H}\cup\tilde{\Xi}}Y) = 0;$$

$$\text{cov}(\tilde{\varepsilon}_j, \tilde{U}) = \text{cov}(Q_{\tilde{H}_{(j)}\cup\tilde{\Xi}}\tilde{\eta}_j, Q_{Y\cup\tilde{\Xi}}X) = 0;$$

$$\text{cov}(\tilde{\Delta}, \tilde{U}) = \text{cov}(Q_{\tilde{H}\cup\tilde{\Xi}}Y, Q_{Y\cup\tilde{\Xi}}X) = 0;$$

$$\text{cov}(\tilde{H}, \tilde{\Delta}) = \text{cov}(\tilde{H}, Q_{\tilde{H}\cup\tilde{\Xi}}Y) = 0 \tag{20}$$

4.2. *Model with a priori information on the variables*

In the case of uncorrelation between pairs of latent variables $\tilde{\eta}_\pi$ and $\tilde{\eta}_\varphi$ or $\tilde{\xi}_\gamma$ and $\tilde{\xi}_\delta$ as defined, we obtain:

$$\eta^0_\pi = Q_{\tilde{\eta}_\varphi} \tilde{\eta}_\pi \quad (\varphi \neq \pi); \qquad \xi^0_\gamma = Q_{\tilde{\xi}_\delta} \tilde{\xi}_\gamma \quad (\gamma \neq \delta) \tag{21}$$

Now, we arrive at new values for $\tilde{E}$, $\tilde{\Delta}$, $\tilde{U}$, Λ_y, Λ_x, B, Γ, according to the previous method.

In order to take into account the uncorrelations between errors in variables $(\tilde{\delta}_{\varphi_\kappa}, \tilde{\delta}_{\varphi_\theta})$ and $(\tilde{u}_{\gamma_\lambda}, \tilde{u}_{\gamma_\mu})$ we get (1):

$$\delta^0_{\varphi_\kappa} = Q_{H^0 \cup \Xi^0 \cup \tilde{\delta}_{\varphi_\theta}} y_{\varphi_\kappa} = Q_{H^0 \cup \Xi^0 \cup y_{\varphi_\theta}} y_{\varphi_\kappa}; \; u^0_{\gamma_\lambda} = Q_{X \cup \Xi^0 \cup \tilde{u}_{\gamma_\mu}} x_{\gamma_\lambda} = Q_{X \cup \Xi^0 \cup x_{\gamma_\mu}} x_{\gamma_\lambda} \tag{22}$$

The model with *a priori* information on the variables contains a further modification of the RCD, which means a new Restricted Regression Component Decomposition (RRCD) of y_{φ_κ}, x_{γ_λ}.

We see that there are no changes in Λ_y, Λ_x, β, Γ (2):

$$b_{(\alpha,\varphi)} = \tilde{\eta}'_\alpha Q_{\tilde{H}_{((\alpha),(\varphi))}} Q_{\tilde{\eta}_\pi} \tilde{\eta}_\varphi / \tilde{\eta}'_\varphi Q_{\tilde{\eta}_\pi} Q_{\tilde{H}_{(\varphi)}} Q_{\tilde{\eta}_\pi} \tilde{\eta}_\varphi = \tilde{\eta}'_\alpha Q_{\tilde{H}_{((\alpha),(\varphi))}} \tilde{\eta}_\varphi / \tilde{\eta}'_\varphi Q_{\tilde{H}_{(\varphi)}} \tilde{\eta}_\varphi \tag{23a}$$

$$\gamma_{(\beta,\gamma)} = \eta'_\beta Q_{\tilde{H}_{(\beta)} \cup \tilde{\Xi}_{(\gamma)}} Q_{\tilde{\xi}_\delta} \tilde{\xi}_\gamma / \tilde{\xi}'_\gamma Q_{\tilde{\xi}_\delta} Q_{\tilde{\Xi}_{(\gamma)}} Q_{\tilde{\xi}_\delta} \tilde{\xi}_\gamma = \eta'_\beta Q_{\tilde{H}_{(\beta)} \cup \tilde{\Xi}_{(\gamma)}} \tilde{\xi}_\gamma / \tilde{\xi}'_\gamma Q_{\tilde{\Xi}_{(\gamma)}} \tilde{\xi}_\gamma \tag{23b}$$

$$l_{x(\varphi_\kappa,\varphi)} = x'_{\varphi_\kappa} Q_{\tilde{H}_{(\varphi)}} Q_{\tilde{\eta}_\pi} \tilde{\eta}_\varphi / \tilde{\eta}'_\varphi Q_{\tilde{\eta}_\pi} Q_{\tilde{H}_{(\varphi)}} Q_{\tilde{\eta}_\pi} \tilde{\eta}_\varphi = x'_{\varphi_\kappa} Q_{\tilde{H}_{(\varphi)}} \tilde{\eta}_\varphi / \tilde{\eta}'_\varphi Q_{\tilde{H}_{(\varphi)}} \tilde{\eta}_\varphi \tag{23c}$$

$$l_{y(\gamma_\lambda,\gamma)} = y'_{\gamma_\lambda} Q_{\tilde{\Xi}_{(\gamma)}} Q_{\tilde{\xi}_\delta} \tilde{\xi}_\gamma / \tilde{\xi}'_\gamma Q_{\tilde{\xi}_\delta} Q_{\tilde{\Xi}_{(\gamma)}} Q_{\tilde{\xi}_\delta} \tilde{\xi}_\gamma = y'_{\gamma_\lambda} Q_{\tilde{\Xi}_{(\gamma)}} \tilde{\xi}_\gamma / \tilde{\xi}'_\gamma Q_{\tilde{\Xi}_{(\gamma)}} \tilde{\xi}_\gamma \tag{23d}$$

---

(1) It is necessary to note that regarding the errors in the variables, defining the errors as residual

$$\text{rank } \Delta = n_y - n_\eta; \qquad \text{rank } U = n_x - n_\xi \tag{N1}$$

Therefore if all the errors Δ or H are mutually non correlated, therefore n_η and n_ξ errors become null.

(2) Only when we use the π-eth equations is this not true because:

$$\tilde{\eta}_\pi Q_{\tilde{H}_{((\pi),(\varphi))}} Q_{\tilde{\eta}_\pi} \tilde{\eta}_\varphi / \tilde{\eta}_\varphi Q_{\tilde{\eta}_\pi} Q_{\tilde{H}_{(\varphi)}} Q_{\tilde{\eta}_\pi} \tilde{\eta}_\varphi \neq \tilde{\eta}'_\pi Q_{H^0_{((\pi),(\varphi))}} \tilde{\eta}_\varphi / \tilde{\eta}'_\varphi Q_{H_{(\varphi)}} \tilde{\eta}_\varphi$$

We have for φ_k-eth equation of the first measurement model, for the γ_λ-eth equation of the second measurement model and for the φ_k-eth equation of the path model the following:

$$V_{Q_{\Xi^0 \cup y_{\varphi_\theta}/H^0} y_{\varphi_k}} = V_{H^0} \oplus V_{y_k / H^0 \cup X \cup y_{\varphi_\theta}}; \; V_{Q_{Y \cup x_{\gamma_\mu}/\Xi^0 \cup x_{\gamma_\lambda}}} = V_{\Xi^0} \oplus V_{x_{\gamma_\lambda} / \Xi^0 \cup Y \cup x_{\gamma_\mu}} \tag{24a}$$

$$V_{\eta^0_\varphi} = V_{Q_{H^0_{(\varphi)}}} \oplus V_{\Xi^0 / H^0_{(\varphi)}} \oplus V_{\eta^0_\varphi / H^0_{(\varphi)} \cup \Xi^0}; \tag{24b}$$

$$Q_{\Xi^0 \cup y^0_{\varphi_\theta}/H^0} y_{\varphi_\kappa} = P_{H^0} y_{\varphi_\kappa} + Q_{H^0 \cup \Xi^0 \cup y_{\varphi_\theta}} y_{\varphi_\kappa}; \; Q_{Y \cup x_{\gamma_\mu}/\Xi^0} x_{\gamma_\lambda} = P_{\Xi^0} x_{\gamma_\lambda} + Q_{Y \cup x_{\gamma_\mu} \cup \Xi^0} x_{\gamma_\lambda} \tag{25a}$$

$$\eta^0_\varphi = P_{H^0_{(\varphi)}} \tilde{\eta}^0_\varphi + P_{\Xi^0 / H^0_{(\varphi)}} \eta^0_\varphi + Q_{H^0_{(\varphi)} \cup \Xi^0} \tilde{\eta}^0_\varphi \tag{25b}$$

with $P_{\Xi^0 / H^0_{(\varphi)}} = Q_{H^0_{(\varphi)}} \Xi^0 \left[\Xi^{0\prime} Q_{H^0_{(\varphi)}} \Xi^0 \right]^{-1} \Xi^{0\prime} Q_{H^0_{(\varphi)}}$

4.3. *Model with restrictions on the parameters*

4.3.1 The RRCD can be further modified by restrictions on the parameters:

$$\beta_{(\mu,\beta)} = \eta^{+\prime}_\mu Q_{H^+_{((\mu),(\beta))}} \eta^+_\beta / \eta^{+\prime}_\beta Q_{H^+_{(\beta)}} \eta^+_\beta = 0$$

$$\gamma_{(\varphi,\delta)} = \eta^{+\prime}_\varphi Q_{H^+_{(\varphi)} \cup \Xi^+_{(\delta)}} \xi^+_\delta / \xi^{+\prime}_\delta Q_{\Xi^+_{(\delta)}} \xi^+_\delta = 0$$

$$l_{y(\alpha,\beta)} = y_\alpha{}' Q_{H^+_{(\beta)}} \eta^+_\beta / \eta^{+\prime}_\beta Q_{H^+_{(\beta)}} \eta^+_\beta = 0$$

$$l_{x(\gamma,\delta)} = x_\gamma{}' Q_{\Xi^+_{(\delta)}} \xi^+_\delta / \xi^{+\prime}_\delta Q_{\Xi^+_{(\delta)}} \xi^+_\delta = 0 \tag{26}$$

with $H^+_{((\mu),(\beta))}$ elements of the H^+ different from $(\eta^+_\mu, \eta^+_\beta)$ and $\Xi^+_{(\delta)}$ elements of the Ξ^+ different from ξ^+_δ, where the latent variables η^+_β, ξ^+_δ are defined in the following way:

$$\eta^{+}_{\beta} = Q_{*\eta_{\mu}\cup *y_{\alpha}}\eta^{0}_{\beta}(*\eta_{\mu} = Q_{H^{0}_{((\mu),(\beta))}}\eta^{0}_{\mu};\ *y_{\alpha} = Q^{0}_{H_{(\beta)}}y_{\alpha})$$

$$\xi^{+}_{\delta} = Q_{*\eta_{\varphi}\cup *x_{\gamma}}\xi^{0}_{\delta}(*\eta_{\varphi} = Q_{H^{0}_{(\varphi)}\cup\Xi^{0}_{(\delta)}}\eta^{0}_{\varphi};\ *x_{\gamma} = Q_{\Xi^{0}_{(\delta)}}x_{\gamma}) \qquad (27)$$

In this way, if there are no other restrictions on the variables H, the restriction of nullity (26) are respected.

If there are restrictions linked to different latent variables H and Ξ we need to set up, starting from the definitions (21)-(27), an iterative process changing H^0 to H^+, Ξ^0 to Ξ^+. This does not produce zero parameters in the first cycle. The procedure converges when, at step n: ${}_{(n-1)}H^+ \cong {}_{n}H^+$, ${}_{(n-1)}\Xi^+ = {}_{n}\Xi^+$ being ${}_{(n-1)}H^+$, ${}_{n}H^+$, ${}_{(n-1)}\Xi^+$, ${}_{n}\Xi^+$ the latent variables defined at step $n-1$ and n of the iterative process.

4.3.2 Finally, we consider the restrictions where some parameters are put equal to one (one for column Λ_y, Λ_x) according to the scale of the variables when they are not standardized and other parameters are constrained. We have from the previous procedure:

$$1_{y(\alpha_1,\alpha)} = a_1;\quad 1_{x(\gamma_1,\gamma)} = b_1;\quad 1_{y(\alpha_\eta,\alpha)} = \alpha_\eta;\quad 1_{y(\gamma_\eta,\gamma)} = \beta_\eta \qquad (28)$$

and putting:

$$\eta^{\diamond}_{\alpha} = \eta^{+}_{\alpha}/a_1;\quad \xi^{\diamond}_{\gamma} = \xi^{+}_{\gamma}/b_1;\quad \xi^{\diamond}_{\gamma} = (\xi^{+}_{\gamma}/\beta_\eta)\,\alpha_\eta \qquad (29)$$

we obtain

$$1^{\diamond}_{x(\alpha_1,\alpha)} = 1;\quad 1^{\diamond}_{x(\beta_\eta,\beta)} = 1;\quad 1^{\diamond}_{y(\alpha_\eta,\alpha)} = 1_{y(\gamma_\eta,\gamma)} \qquad (30)$$

The RRCD is analogous to (24), (25), with the latent variables defined as in (27).

5. Examples

In order to apply the previous method, we present seven examples (Table 1); in Table 2, there are results reached by means of the method discussed in previous sections. The equations (3) and (4), the properties, and the restrictions of the model are satisfied in every example.

1. Hodge and Treiman (1968) studied the causal linear relation among the latent variable η "social partecipation" (with the indicators y_1 church attendance, y_2 memberships, y_3 friends seen) and the indicators inherent social status [x_1 income, x_2 occupation, x_3 education]. Jöreskog and Sörbom (1981) studied the problem by means of the Lisrel model.

2. Wheaton and others (1977) analyzed the linear relation between the latent variable "stability of alienation" η (with the indicators y_1 Anomia 67, y_2 Powerlessness 67, y_3 Anomia 71, y_4 Powerlessness 71) and a latent variable "background situation" ξ (with the indicators x_1 education and x_2 Duncan socioeconomic index). Jöreskog and Sörbom (1981) studied the problem by means of the Lisrel model.

3. Warren, White and Fuller (1974) studied the role of the behaviour of farmers measured by y_1 "his role performance" and by x_1 "knowledge of economic phases", x_2 "value orientation", x_3 "role satisfaction", x_4 "past training". Jöreskog and Sörbom (1981) studied the problem by means of the Lisrel model.

4. Calsyn and Kenny (1977) presented a model inherent to the linear relations between the latent variables η_1 "ability" and η_2 "aspiration" with the indicators y_1 "self concept of ability", y_2 "parental evaluation", y_3 "teacher evaluation", y_4 "friend's evaluation", y_5 "educational aspiration", y_6 "college plans". Jöreskog and Sörbom (1981) analyzed the problem by means of the Lisrel model(3).

5. Jöreskog and Sörbom (1981) proposed an hypothetical model where the latent variables η_1, η_2 (with 4 indicators y_1, y_2, y_3, y_4) are linearly connected with the latent variables ξ_1, ξ_2, ξ_3 (with 7 indicators $x_1, x_2, x_3, x_4, x_5, x_6, x_7$).

6. Here another hypothetical model is proposed where the latent variables η_1, η_2, η_3 (with the observed indicators $y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8$) are linearly connected with the latent variables ξ_1, ξ_2, ξ_3 (with the observed indicators x_1, x_2, x_3, x_4, x_5) (4).

(3) Unlike them we also calculate the matrix B and the errors E

(4) The solution can be reached by means of a MATLAB program prepared by the authors. They are available at their e-mail addresses:
khaagen@gelso.unitn.it; vittadin@imiucca.csi.unimi.it.

TABLE I

	Y	X	Λ_y	Λ_x	B	Γ	H	Ξ	E	Δ	U
I	y_1, y_2, y_3	x_1, x_2, x_3	$\begin{pmatrix}\lambda_{11}\\ \lambda_{21}\\ \lambda_{31}\end{pmatrix}$	—	$(\beta_1, \beta_2, \beta_3)$	—	η_1	$X=\Xi$	ε_1	$\delta_1, \delta_2, \delta_3$ correlated	—
II	y_1, y_2, y_3, y_4	x_1, x_2	$\begin{pmatrix}1 & 0\\ \lambda_{12} & 0\\ 0 & 1\\ 0 & \lambda_{42}\end{pmatrix}$	$\begin{pmatrix}1\\ \lambda_{21}\end{pmatrix}$	$\begin{pmatrix}0 & \beta_{12}\\ \beta_{21} & 0\end{pmatrix}$	$\begin{pmatrix}\gamma_{11}\\ \gamma_{21}\end{pmatrix}$	η_1, η_2 correlated	ξ_1	$\varepsilon_1, \varepsilon_2$ uncorrelated	δ_1, δ_3 correlated; the others uncorrelated	u_1, u_2 uncorrelated
III	y_1	x_1, x_2, x_3, x_4	—	—	—	$(\gamma_1\ \gamma_2\ \gamma_3\ \gamma_4)$	$Y=H$	$X=\Xi$	ε_1	—	—
IV	$y_1, y_2, y_3, y_4, y_5, y_6$		$\begin{pmatrix}\lambda_{11} & 0\\ \lambda_{21} & 0\\ \lambda_{31} & 0\\ \lambda_{41} & 0\\ 0 & \lambda_{52}\\ 0 & \lambda_{62}\end{pmatrix}$	—	$\begin{pmatrix}0 & \beta_{12}\\ \beta_{21} & 0\end{pmatrix}$	—	η_1, η_2 correlated	—	$\varepsilon_1, \varepsilon_2$ uncorrelated	$\delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6$ uncorrelated	—

TABLE I (cont.)

	Y	X	Λ_y	Λ_x	B	Γ	H	Ξ	E	Δ	U
V	y_1, y_2, y_3, y_4	$x_1, x_2, x_3, x_4, x_5, x_6, x_7$	$\begin{pmatrix}1 & 0\\ \lambda_{21} & 0\\ 0 & 1\\ 0 & \lambda_{42}\end{pmatrix}$	$\begin{pmatrix}1 & 0 & 0\\ \lambda^{*}_{21} & 0 & 0\\ \lambda_{31} & \lambda_{32} & 0\\ 0 & 1 & 0\\ 0 & \lambda_{52} & 0\\ 0 & 0 & 1\\ 0 & 0 & \lambda_{73}\end{pmatrix}$	$\begin{pmatrix}0 & \beta_{12}\\ \beta_{21} & 0\end{pmatrix}$	$\begin{pmatrix}\gamma_{11} & \gamma_{12} & 0\\ \gamma_{21} & 0 & \gamma_{23}\end{pmatrix}$	η_1, η_2 correlated	ξ_1, ξ_2, ξ_3 correlated	$\varepsilon_1, \varepsilon_2$ uncorrelated	$\delta_1, \delta_2, \delta_3, \delta_4$ uncorrelated	$u_1, u_2, u_3, u_4, u_5, u_6, u_7$ uncorrelated
VI	$y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8$	x_1, x_2, x_3, x_4, x_5	$\begin{pmatrix}\lambda_{11} & 0 & 0\\ \lambda_{21} & 0 & 0\\ \lambda_{31} & 0 & 0\\ \lambda_{41} & \lambda_{42} & 0\\ \lambda_{51} & \lambda_{52} & 0\\ \lambda_{61} & \lambda_{62} & \lambda_{63}\\ 0 & \lambda_{72} & \lambda_{73}\\ 0 & 0 & \lambda_{83}\end{pmatrix}$	$\begin{pmatrix}\lambda_{11} & 0 & 0\\ \lambda_{21} & \lambda_{22} & 0\\ \lambda_{31} & \lambda_{32} & \lambda_{33}\\ 0 & \lambda^{*}_{42} & \lambda_{43}\\ 0 & 0 & \lambda_{53}\end{pmatrix}$	$\begin{pmatrix}0 & 0 & \beta_{13}\\ 0 & 0 & \beta_{23}\\ \beta_{31} & \beta_{32} & 0\end{pmatrix}$	$\begin{pmatrix}0 & \gamma_{12} & \gamma_{13}\\ \gamma_{21} & 0 & \gamma_{23}\\ \gamma_{31} & \gamma_{32} & 0\end{pmatrix}$	η_1, η_2, η_3; η_1 uncorrelated with η_2	ξ_1, ξ_2, ξ_3; ξ_1 uncorrelated with ξ_2	$\varepsilon_1, \varepsilon_2, \varepsilon_3$ uncorrelated	$\delta_1, \delta_2, \delta_3$ correlated; $\delta_5, \delta_6, \delta_7, \delta_8$ correlated; δ_4 uncorrelated with any other error; $\delta_1, \delta_2, \delta_3$ uncorrelated with $\delta_5, \delta_6, \delta_7, \delta_8$	u_1, u_4, uncorrelated with any other error; u_2 correlated only with u_3; u_5 correlated only with u_3.

TABLE 2

	Variance-covariance matrix	Λ_Y	Λ_X	B	Γ
I	1 3040 1 3050 3440 1 1000 1560 1580 1 2840 1920 3240 3600 1 1760 1360 2260 2100 2650 1	0.5608 0.5952 0.5857	—	(0.0591 0.3884 0.1953)	
II	11.834 6.947 9.364 6.819 5.091 12.532 4.783 5.028 7.495 9.986 −3.839 −3.889 −3.841 −3.625 9.610 −21.899 −18.831 −21.748 −18.775 35.522 450.288	1 0 0.8403 0 0 1 0 0.8403	1 12.4442	0 0.5802 0.6202 0	−0.291 −0.258
III	0.0209 0.0177 0.0520 0.0245 0.0280 0.1212 0.0046 0.0044 −0.0063 0.0901 0.0187 0.0192 0.0353 −0.0066 0.0946	—	—	—	(0.2305 0.1197 0.0562 0.1101)
IV	1 0.73 1 0.70 0.68 1 0.58 0.61 0.57 1 0.46 0.43 0.40 0.37 1 0.56 0.52 0.48 0.41 0.72 1	−0.2407 0 −0.4871 0 0.8361 0 0.1306 0 0 0.7098 0 −0.7044	—	0 0.0399 1.395 0	—

TABLE 2 (cont.)

	Variance-covariance matrix	Λ_Y	Λ_X	B	Γ
V	3.815 3.195 3.068 3.890 3.508 5.773 4.259 3.842 6.182 6.973 0.764 0.690 0.176 0.193 1.374 0.994 0.896 0.229 0.251 1.298 1.952 1.690 1.525 0.347 0.380 1.739 2.260 4.224 0.835 0.753 0.157 0.172 0.700 0.909 2.069 1.471 0.917 0.827 0.173 0.189 0.768 0.998 2.272 1.317 1.822 1.259 1.136 1.824 1.997 0.601 0.781 0.902 0.300 0.330 1.675 1.763 1.591 2.554 2.797 0.841 1.094 1.262 0.421 0.462 1.958 2.914	1 0 0.9295 0 0 1 0 −1.5714	1 0 0 −0.8069 0 0 0.0788 −0.1739 0 0 1 0 0 0.5611 0 0 0 1 0 0 −0.8433	0 −0.0305 −0.0016 0	−0.0625 0.0313 0 −0.0002 0 −0.0005
VI	0.0891 −0.0096 0.0951 −0.0162 −0.0017 0.0816 −0.0043 −0.0072 −0.0079 0.0755 0.0202 −0.0126 −0.0301 −0.0103 0.0844 −0.0068 0.0037 0.0037 0.0167 −0.0163 0.0678 0.0084 −0.0001 0.0263 0.0175 −0.0198 0.0081 −0.0972 −0.0082 0.0027 −0.0024 0.0049 −0.0080 −0.0142 −0.0142 0.0724 0.0143 0.0125 −0.0085 −0.0048 0.0171 −0.0018 0.0107 −0.0057 0.0493 0.0293 −0.0216 0.0089 0.0032 −0.0060 0.0112 −0.0056 0.0137 −0.0130 0.0834 0.0058 0.0060 −0.0064 −0.0062 0.0294 −0.0203 0.0042 0.0057 0.0132 −0.0291 0.0788 0.0122 −0.0124 −0.0079 0.0378 −0.0255 0.0035 0.0228 −0.0143 −0.0024 0.0029 0.0077 0.0931 −0.0093 −0.0177 −0.0005 0.0171 0.0057 −0.0032 0.0156 −0.0181 −0.0001 0.0086 0.0065 0.0028 0.0767	−0.6659 0 0 −0.3077 0 0 −0.6811 0 0 0.1227 1.0460 0 −0.0671 0.1949 0 −0.0010 0.7465 −0.3127 0 0.8164 0.2761 0 0 0.2675	1.0165 0 0 0.0896 0.6452 0 −0.0546 0.6998 0.4009 0 0.3065 −0.9160 0 0 0.0580	0 0 −0.0658 0 0 0.0038 −0.1715 0.0149 0	0 −0.2903 −0.2874 0.1372 0 −0.1442 0.0317 0.2945 0

TABLE 2 (cont.)

	Σ_H	Σ_Ξ	Σ_E	Σ_Δ	Σ_U
I	1.6357	—	1.3817	$\begin{pmatrix} 0.4826 & -0.2284 & -0.2340 \\ -0.5584 & 0.4296 & -0.2105 \\ -0.2340 & -0.2105 & 0.4343 \end{pmatrix}$	—
II	$\begin{pmatrix} 10.3430 & 6.4142 \\ 6.4143 & 11.0555 \end{pmatrix}$	2.885	$\begin{pmatrix} 6.400 & 0 \\ 0 & 6.905 \end{pmatrix}$	$\begin{pmatrix} 1.4917 & 0 & 0.4098 & 0 \\ 0 & 0.0221 & 0 & 0 \\ 0.4098 & 0 & 7.013 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} \text{[illegible]} & 0 \\ \text{[illegible]} & 0 \end{pmatrix}$
III	—	—	0.0116	—	—
IV	$\begin{pmatrix} 0.3239 & 0.0453 \\ 0.0453 & 1.1345 \end{pmatrix}$	—	$\begin{pmatrix} 1.1282 & 0 \\ 0 & 0.3221 \end{pmatrix}$	$\begin{pmatrix} 0.9723 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.4327 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.0066 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.0366 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$	—

TABLE 2 (cont.)

	Σ_H	Σ_Ξ	Σ_E	Σ_Δ	Σ_U
V	$\begin{pmatrix} 0.9861 & -0.0015 \\ -0.0015 & 0.0501 \end{pmatrix}$	$\begin{pmatrix} 0.1752 & 0.0043 & 0.0007 \\ 0.0043 & 0.2515 & 0.0003 \\ 0.0007 & 0.0003 & 0.1676 \end{pmatrix}$	$\begin{pmatrix} & 0.9850 \end{pmatrix}$	$\begin{pmatrix} 2.8256 & 0 & 0 & 0 \\ 0 & 0.3793 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0.7727 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.8135 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.3290 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.0216 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$
VI	$\begin{pmatrix} 0.0577 & 0 & -0.0031 \\ 0 & 0.0379 & -0.0035 \\ -0.0031 & -0.0035 & -0.0771 \end{pmatrix}$	$\begin{pmatrix} 0.0425 & 0 & 0.0072 \\ 0 & 0.0516 & -0.0085 \\ 0.0072 & -0.0085 & 0.0891 \end{pmatrix}$	$\begin{pmatrix} 0.0402 & 0 & 0 \\ 0 & & 0.0311 \\ & & 0 \end{pmatrix}$	$\begin{pmatrix} 0.0654 & -0.079 & -0.0104 & 0 & 0 & 0 & 0 & 0 \\ -0.0975 & 0.0701 & 0.0226 & 0 & 0 & 0 & 0 & 0 \\ -0.0104 & 0.0026 & 0.0488 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.573 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.0709 & -0.0112 & -0.0277 & -0.096 \\ 0 & 0 & 0 & 0 & -0.0111 & 0.0496 & 0.0170 & -0.0187 \\ 0 & 0 & 0 & 0 & -0.0277 & 0.0170 & 0.0891 & -0.0138 \\ 0 & 0 & 0 & 0 & -0.0096 & 0.0187 & -0.0138 & 0.0665 \end{pmatrix}$	$\begin{pmatrix} 0.3000 & 0 & 0 & 0 & 0 \\ 0 & 0.0499 & -0.0232 & 0 & 0 \\ 0 & -0.0232 & 0.0469 & 0 & -0.018 \\ 0 & 0 & 0 & 0.05010 & 0 \\ 0 & 0 & -0.018 & 0 & 0.044 \end{pmatrix}$

REFERENCES

APEL, H. and WOLD, H. (1982) Soft modeling with latent variables in two or more dimensions: PLS estimation and testing for predictive relevance, in JORESKOG K.G. and WOLD H. (eds.), *System under indirect observation: causality, structure, prediction*, North-Holland, Amsterdam, 2, 209-247.

BEKKER, P.A. and POLLOCK, D.S.G. (1985) Identification of linear stochastic models with covariance restrictions, *Journal of Econometrics*, 31, 179-206.

BEKKER, P.A. (1989) Identification in restricted factor models and the evaluation of rank conditions, *Journal of Econometrics*, 41, 5-15.

BENTLER, P.M. (1976) Multistructure statistical model applied to factor analysis, *Multivariate behavioral research*, 11, 3-25.

BENTLER, P.M. (1982) Linear systems with multiple levels and types of latent variables, in JORESKOG K.G. and WOLD H. (eds.), *Systems under indirect observation: causality, structure, prediction*, North-Holland, Amsterdam, 1, 101-130.

BENTLER, P.M. (1983) Simultaneous equation systems as moment structure models, *Journal of Econometrics*, 22, 13-42.

BENTLER, P.M. and WEEKS, D.G. (1980) Linear structural equations with latent variables, *Psychometrika*, 45, 3, 289-308.

DEMPSTER, A.P. (1971) An overview of multivariate data analysis, *Journal of Multivariate Analysis*, 1, 316-346.

DIJKSTRA, T. (1983) Some comments on maximum likelihood and partial least squares methods, *Journal of Econometrics*, 22, 67-90.

DUNCAN, O.D., HALLER, A.O. and PORTES, A. (1968) Peer influence on aspiration: a reinterpretation, *American Journal of Sociology*, 74, 119-137.

EVERITT, B.S. (1984) *An introduction to latent variable models*, Chapman and Hall (Eds.), London.

FISHER, F.M. (1966) *The identification problem in econometrics*, McGraw Hill, New York.

GERACI, V.J. (1976) Identification of simultaneous equation models with measurement error, *Journal of Econometrics*, 4, 263-283.

GUTTMAN, L. (1955) The determinacy of factor score matrices with implications for five other basic problems of common factor theory, *The British Journal of Statistical Psychology*, 8, L, 65-81.

HAAGEN, K. (1986) The indeterminacy problem in factor analysis and its consequences for prediction, *Österreichische Zeitschrift für Statistik und informatik*, 4, 197-205.

HAAGEN, K. (1987) On the use of factor scores for prediction, *Biometrical Journal*, 29, 7, 835-847.

HAAGEN, K. (1991) Necessary and sufficient conditions for vanishing the arbitrariness in a common factor model, *The Statistician*, 40, 401-408.

HAAGEN, K. and VITTADINI, G. (1991) Regression component decomposition in structural analysis, *Communication in statistics - Theory and methods*, New York, 20, 4, 1153-1161.

HODGE, R.W. and TREIMAN, D.J. (1968) Social participation and social status, *American sociological review*, 33, 722-739.

JÖRESKOG, K.G. (1970) A general method for analysis of covariance structures, *Biometrika*, 57, 2, 239-251.

JÖRESKOG, K.G. (1971) Simultaneous factor analysis in several populations, *Psychometrika*, 36, 409-426.

JÖRESKOG, K.G. (1973a) A general method for estimating a linear structural equation system, in GOLDBERGER A.S. and DUNCAN O.D. (eds.), *Structural equation models in the social sciences*, Seminar Press, New York, 85-112.

JÖRESKOG, K.G. (1973b) Analysis of covariance structures, in KRISHNAIAH P.R. (ed.), *Multivariate analysis*, 3 Academic Press, New York, 263-285.

JÖRESKOG, K.G. (1974) Analyzing psychological data by structural analysis of covariance matrices, in ATKINSON, R.C., KRANTZ, D.H., LUCE, R.D. and SUPPES P. (eds.), *Contemporary developments in mathematical psychology*, Freeman, W.H., San Francisco, 2, 1-56.

JÖRESKOG, K.G. (1977) Structural equation models in the social sciences: specification, estimation and testing, in: *Application of statistics*, Krishnaiah P.R., Amsterdam, North-Holland, 265-287.

JÖRESKOG, K.G. (1978) Structural analysis of covariance and correlation matrices, *Psycometrika*, 43, 443-477.

JÖRESKOG, K.G. (1979a) Basic ideas of factor and component analysis, in JÖRESKOG K.G. and SORBOM D. (eds.), *Advances in factor analysis and structural equation models*, (eds.) Abt Books, Cambridge, Massachussets, 5-20.

JÖRESKOG, K.G. (1979b) Statistical estimation of structural models in longitudinal-development investigations, in NESSELROADE J.R. and BALTES, P.B. (eds.), *Longitudinal research in the study of behavior and development*, Academic Press, New York, 303-351.

JÖRESKOG, K.G. (1981a) Basic issues in the applications of LISREL, *Data communications computer data analysis*, 1, 1-6.

JÖRESKOG, K.G. (1981b) Analysis of covariance structures, *Scandinavian Journal of Statistics*, 8, 65-92.

JÖRESKOG, K.G. and SÖRBOM, D. (1981) *LISREL VI*, Chicago: International Educational Services.

JÖRESKOG, K.G. (1982) The LISREL approach to casual model-building in the social sciences, in JÖRESKOG, K.G. and WOLD H. (eds.), *Systems under indirect observation: causality, structure, prediction*, North-Holland, Amsterdam, 1, 81-99.

KMENTA, J. (1991) Latent variables in econometrics, *Statistica Neerlandica*, 45, 2, 73-83.

MCDONALD, R.P. (1978) A simple comprehensive model for the analysis of covariance structures, *British Journal of Mathematical and Statistical Psychology*, 31, 59-72.

MULAIK, S.A. (1972) *The foundations of factor analysis*, New York, McGraw Hill.

NOONAN, R. and WOLD, H. (1982) PLS Path modeling with indirectly observed variables, in JÖRESKOG, K.G. and WOLD H. (eds.), *System under indirect observation: causality, structure, prediction*, North-Holland, Amsterdam, 2, 75-94.

ROTHENBERG, T.J. (1971) Identification in parametric models, *Econometrica*, 39, 3, 577-591.

SCHÖNEMANN, P.H. and HAAGEN, K. (1987) On the use of factor scores for prediction, *Biometrical Journal*, 29, 7, 835-847.

SCHÖNEMANN, P.H. and WANG, M.M. (1972) Some new results on factor indeterminancy, *Psychometrica*, 37, 1, 61-91.

SCHÖNEMANN, P.H. and STEIGER, J.H. (1976) Regression component analysis, *British Journal of mathematical and statistical psychology*, 29, 175-189.

SCHÖNEMANN, P.H. and STEIGER, J.H. (1978) On the validity of indeterminate factor scores, *Bulletin Psyconomic Soc.*, 12, 287-290.

SCHNEEWEIß, H. (1991) Models with latent variables: LISREL versus PLS, *Statistica Neerlandica*, 45, 2, 145-171.

SHAPIRO, A. (1985) Identificability of factor analysis: some results and open problems, *Linear algebra and its applications*, 70, 1-7.

TAKEUCHI, K., YANAI, H. and MUKHERJEE, B.N. (1982) *The foundations of multivariate analysis*, Wiley Eastern Limited, New Delhi.

VITTADINI, G. (1988) On the validity of the indeterminate latent variables in the LISREL model, *Communications in statistics - Theory and methods*, New York, 17, 3, 861-874.

VITTADINI, G. (1989) Indeterminacy problems in the Lisrel Model, in *Multivariate Behavioral Research*, Forth Worth (Texas), 24, 4, 397-414.

VITTADINI, G. (1992) Un confronto fra un modello di analisi causale con variabili latenti e metodi alternativi, *Statistica*, LII, 3, 379-396.

WARREN, R.D., WHITE, J.K. and FULLER, W.A. (1974) An error in variables analysis of managerial role performance, *Journal of the American Statistical Association*, 69, 886-893.

WEGGE, L.L. (1965) Identifiability criteria for a System of equation as a whole, *The Australian Journal of Statistics*, 7(3), 67-77.

WEGGE, L.L. (1991) Identification with latent variables, *Statistical Neerlandica*, 45, 2, 121-143.

WEGGE, L.L. (1996) Local identifiability of the factor analysis and measurement error model parameter, *Journal of Econometrics*, 70, 351-382.

WHEATON, B., MUTHEN, B., ALWIN, D. and SUMMERS, G. (1977) Assessing reliability and stability in panel models, in HEISE, D.R. (ed.) *Sociological methodology 1977*, San Francisco: Jossey-Bass, 84-136.

WILLIAMS, J.S. (1978) A definition for the common factor analysis model and the elimination of problems of factor score indeterminacy, *Psychometrica*, 43, 3, 293-306.

WOLD, H. (1980) Model construction and evaluation when theoretical knowledge is scarce: Theory and application of partial least squares, in KMENTA, J. and RAMSEY, J.B. (eds.), *Model evaluation in econometrics*, Academic Press, New York, 47-74.

WOLD, H. (1982) Soft modeling: the basic design and some extensions, in JÖRESKOG, K.G. and WOLD, H., *Systems under indirect observation: causality, structure, prediction*, North-Holland, Amsterdam, 2, 1-54.

Restricted regression component decomposition

SUMMARY

The structural models do not have unique solutions under general conditions. The alternative proposals (Partial Least Squares, Regression Component Decompositions) leave some unsolved problems where there are some *a priori* information and some restrictions on the model. A method is proposed, Restricted Regression Component Decomposition, which defines latent variables and derives relations from covariance structures, in a flexible way respect to *a priori* information and restrictions.

Decomposizione in componenti di regressione vincolate

RIASSUNTO

I modelli strutturali non hanno, sotto condizioni generali, soluzioni uniche. Le proposte alternative (Minimi Quadrati Parziali, Decomposizione in Componenti di Regressione) lasciano irrisolti alcuni problemi quando vi sono informazioni a priori e restrizioni sul modello. È proposto un metodo, denominato Decomposizione in Componenti di Regressione Vincolate, che definisce variabili latenti e deriva relazioni dalla struttura delle covarianze in modo flessibile rispetto ad informazioni a priori e restrizioni.

KEY WORDS

Structural model; identification problem; regression component decomposition; restricted regression component decomposition.

[Manuscript received November, 1997; final version received March, 1998]