A New Version of the Structural Dynamic Model with Unique Latent Scores

Simona Caterina Minotti[1] and Giorgio Vittadini[2]

[1] Dipartimento di Scienze Economiche e Sociali,
Università Cattolica di Piacenza, Italy
simona.minotti@unicatt.it

[2] Dipartimento di Statistica,
Università Milano Bicocca, Italy
giorgio.vittadini@unimib.it

Abstract. The indeterminacy of the Structural Models, i.e. the arbitrariness of latent scores, due to the factorial nature of the measurement models, is, in the dynamic context, more problematic. We propose an alternative formulation of the Structural Dynamic Model, based on the Replicated Common Factor Model (Haagen e Oberhofer, 1999), where latent scores are no more indeterminate.

1 Introduction

It is well known that the causality principle of the Factor Analysis Model (FA) (i.e. to express the indicators as a function of the latent variables) leads to indeterminacy of latent scores (Guttman (1955)), with important consequences on the classificatory validity of the latent variables (Schoenemann and Haagen (1987); Haagen (1991)). The same problem arises in the Structural Equation Models with reflective blocks (SEM) (Vittadini (1989)) and also in the Structural Dynamic Models with latent variables (SDL) (Haagen and Vittadini (1994)), because both use FA as a measurement model. Moreover Haagen and Vittadini (1994) proved that the dynamics increases the indeterminacy. The SDL will be presented in Section 2, while the problem of indeterminacy in SDL will be discussed in Section 3.

It follows that only the introduction of alternative models to the FA as measurement models allows us to definitely overcome indeterminacy. To avoid arbitrariness of common factors, some authors have proposed an estimation procedure which inverts the natural relationship among variables, expressing the latent variables as a linear combination of the indicators. For example, Schoenemann and Steiger (1976) proposed the Regression Component Decomposition method (RCD), successively extended to SEM by Haagen and Vittadini (1991) and applied to some different contexts (Vittadini (1999)). This proposal, however, can not be considered a model, because the solutions can not be interpreted as causes of the indicators. Instead Haagen and Oberhofer (1999) introduced an alternative model to the FA, entitled the Replicated Common Factor Model (RCFM), (successively extended to SEM

by Vittadini and Haagen (2002)), which solves the indeterminacy of common factors asymptotically.
Nevertheless, no proposal has been made in the dynamic context, except by Minotti (2002), who introduced a correction of the Kalman filter by means of RCD, providing, at every time interval t, (t=1,...,T), unique values for the latent variables. This new method will be described in Section 3. However, as illustrated previously, the proposal is not a model. Thus, the problem of indeterminacy in the SDL can not be considered definitively overcome.
As an alternative, which overcomes the indeterminacy of latent scores asymptotically, we propose a new version of the SDL based on the RCFM. The extension of the RCFM to the dynamic context will be presented in Section 4. The new model has been applied to an experiment of the Department of Physics, University of Milan. The application will be described in Section 5. Some conclusions will be given in Section 6.
Section 1 is to be attributed to Vittadini, as well as the supervision of the paper; Sections 2-6 were developed by Minotti.

2 The SDL

The SDL, introduced by Otter (1985) as a dynamic generalization of SEM, is the stationary version of the Stochastic Linear State Space Model from Systems Engineering, i.e. it is a linear model with latent variables, where the observations are represented by a single multivariate time series. The SDL consists of a transition equation, which describes the temporal relationships among the latent variables, and a measurement equation, that relates the latent variables to the observed variables:

$$\mathbf{A}_0\boldsymbol{\xi}_t = \mathbf{A}_1\boldsymbol{\xi}_{t-1} + \overline{\mathbf{B}}\mathbf{u}_t + \overline{\mathbf{w}}_t, \qquad \boldsymbol{\xi}_0 \sim N(\boldsymbol{\mu}, \mathbf{V}_{\boldsymbol{\xi}_0}), \qquad t = 1, ..., T \tag{1}$$

$$\mathbf{z}_t = \mathbf{C}\boldsymbol{\xi}_t + \mathbf{D}\mathbf{u}_t + \mathbf{v}_t, \qquad t = 1, ..., T \tag{2}$$

where $\boldsymbol{\xi}_t = [\boldsymbol{\eta}_t, \boldsymbol{\phi}_t]$ are respectively m_1 endogenous and m_2 exogenous latent variables distributed as normal random variables with finite covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\xi}_t}$ ($\boldsymbol{\xi}_0 \sim N(\boldsymbol{\mu}, \mathbf{V}_{\boldsymbol{\xi}_0})$); $\overline{\mathbf{w}}_t$ and $\mathbf{v}_t = [\boldsymbol{\epsilon}_t, \boldsymbol{\delta}_t]$ are respectively $m = m_1 + m_2$ and $p = p_1 + p_2$ latent errors, normally distributed and mutually non correlated (for every t), with null expected value and time-invariant covariance matrices $\boldsymbol{\Sigma}_{\overline{\mathbf{W}}}$ and $\boldsymbol{\Sigma}_{\mathbf{V}}$; $\mathbf{u}_t$ is a vector of q deterministic inputs; $\mathbf{A}_0, \mathbf{A}_1, \mathbf{B}, \mathbf{C}$ and $\mathbf{D}$ are time-invariant matrices of respectively $(m \times m)$, $(m \times m)$, $(m \times q)$, $(p \times m)$ and $(p \times q)$ parameters (with invertible $\mathbf{A}_0$); $\mathbf{z}_t = [\mathbf{y}_t, \mathbf{x}_t]$ are p indicators respectively of the $\boldsymbol{\eta}_t$ and the $\boldsymbol{\phi}_t$, with covariance matrix $\boldsymbol{\Sigma}_{\mathbf{Z}_t}$.
The reduced form of the SDL is obtained pre-multiplying (1) by $\mathbf{A}_0^{-1}$, i.e.:

$$\boldsymbol{\xi}_t = \mathbf{A}_0^{-1}\mathbf{A}_1\boldsymbol{\xi}_{t-1} + \mathbf{A}_0^{-1}\overline{\mathbf{B}}\mathbf{u}_t + \mathbf{A}_0^{-1}\overline{\mathbf{w}}_t, \tag{3}$$

expressed in a compact notation as

$$\boldsymbol{\xi}_t = \mathbf{A}\boldsymbol{\xi}_{t-1} + \mathbf{B}\mathbf{u}_t + \mathbf{w}_t, \tag{4}$$

where $\mathbf{w_t}$ is a vector of m latent errors with null expected value and time-invariant covariance matrices $\mathbf{\Sigma_W} = \mathbf{A}_0^{-1}\mathbf{\Sigma}_{\overline{\mathbf{W}}}(\mathbf{A}_0^{-1})'$.
In the following, the deterministic inputs $\mathbf{u}_t$ will be omitted, which is not an essential restriction, given that they can always be included among the observable variables.
The parameter identifiability of the SDL has been extensively studied by many authors, who proposed some conditions for the local identifiability (Bordignon and Trivellato (1992); Otter (1992)). The estimation of the parameters and the latent scores can be obtained by means of a recursive procedure (as illustrated in Haagen and Vittadini (1994)).

3 The indeterminacy in the SDL and the first solution

Haagen and Vittadini (1994) discussed the indeterminacy of the SDL solution by using the Guttman's result (1955), introduced for the FA. They demonstrated that, if the following equality holds

$$\mathbf{\Sigma}_{\mathbf{Z}_t} = \mathbf{C}\mathbf{\Sigma}_{\boldsymbol{\xi}_t}\mathbf{C}' + \mathbf{\Sigma}_{\mathbf{V}_t}, \tag{5}$$

there exist arbitrary vectors $\boldsymbol{\xi}_t$ which satisfy (2).
These solutions, called "true solutions", have the following structure:

$$\boldsymbol{\xi}_t = \widehat{\boldsymbol{\xi}}_{t|t-1} + \mathbf{K}_t[\mathbf{z}_t - \widehat{\mathbf{z}}_t] + \mathbf{P}_t\boldsymbol{\omega}_t = \widehat{\boldsymbol{\xi}}_{t|t} + \mathbf{P}_t\boldsymbol{\omega}_t, \tag{6}$$

where $\widehat{\boldsymbol{\xi}}_{t|t-1}$ is the efficient predictor for $\boldsymbol{\xi}_t$, based linearly on $\{\mathbf{z}_1, ..., \mathbf{z}_{t-1}\}$, $\widehat{\boldsymbol{\xi}}_{t|t}$ is the Kalman estimator of $\boldsymbol{\xi}_t$, updated on the basis of the last observations $\mathbf{z}_t$, $\mathbf{K}_t = \mathbf{\Sigma}_{\widehat{\boldsymbol{\xi}}_{t|t-1}}\mathbf{C}'\mathbf{\Sigma}_{\mathbf{Z}_t}^{-1}$ is the Kalman gain, $\widehat{\mathbf{z}}_t = \mathbf{C}\widehat{\boldsymbol{\xi}}_{t|t-1}$ is the forecasting of $\mathbf{z}_t$ by (2), $\mathbf{P}_t = (\mathbf{I} - \mathbf{K}_t\mathbf{C})\mathbf{F}_t$ with $\mathbf{F}_t\mathbf{F}_t' = \mathbf{\Sigma}_{\widehat{\boldsymbol{\xi}}_t}$, $\boldsymbol{\omega}_t$ is an arbitrary vector of m variables and $E[\boldsymbol{\omega}_t] = \mathbf{0}$, $\mathbf{\Sigma}_{\boldsymbol{\omega}_t,\mathbf{Z}_t} = \mathbf{0}$, $\mathbf{\Sigma}_{\boldsymbol{\omega}_t} = \mathbf{I} - \mathbf{C}'\mathbf{\Sigma}_{\mathbf{Z}_t}^{-1}\mathbf{C}$.
It follows that, as in the static context (FA and SEM), the latent scores obtained by means of the Kalman filter are not unique, due to the arbitrary term $\mathbf{P}_t\boldsymbol{\omega}_t$. Moreover Haagen and Vittadini (1994) demonstrated that in the SDL indeterminacy increases, due to the recursive procedure of Kalman filter, which spreads indeterminacy in time (the indeterminacy of $\widehat{\boldsymbol{\xi}}_t$ influences $\widehat{\boldsymbol{\xi}}_{t+1}$, $\widehat{\boldsymbol{\xi}}_{t+2}$, ...). In addition to this, Minotti (2002) observes that in the dynamic context defining $\widehat{\mathbf{z}}_t$ by (2) makes the situation worse.
As a first attempt of solution, Minotti (2002) proposed a correction of the Kalman filter by means of RCD, described in detail in the following.
Referring to the formulation of Haagen and Vittadini (1991), the RCD provides at each date t the following decomposition of $\mathbf{x}_t$:

$$\mathbf{x}_t = \widetilde{\mathbf{C}}_t\widetilde{\boldsymbol{\phi}}_t + (\mathbf{x}_t - \widetilde{\mathbf{C}}_t\widetilde{\boldsymbol{\phi}}_t), \tag{7}$$

which leads to a definition of the "latent" variables $\widetilde{\boldsymbol{\phi}}_t$, called components, as a linear combination of the observed variables $\mathbf{x}_t$

$$\widetilde{\boldsymbol{\phi}}_t = \mathbf{L}_{\boldsymbol{\Phi}_t}'\mathbf{x}_t, \tag{8}$$

where $\mathbf{L}_{\boldsymbol{\Phi}_t} = \boldsymbol{\Sigma}_{\mathbf{X}_t}^{-1}\widetilde{\mathbf{C}}_t(\widetilde{\mathbf{C}}_t'\boldsymbol{\Sigma}_{\mathbf{X}_t}^{-1}\widetilde{\mathbf{C}}_t)^{-1}$ and $\widetilde{\mathbf{C}}_t$ can be calculated as a factor loading matrix by means of a factor extraction method (Schoenemann and Steiger (1976)).
By means of an analogous decomposition of $\mathbf{y}_t$, we obtain $\mathbf{L}_{\boldsymbol{\eta}_t}$. The matrix

$$\mathbf{L}_t = \begin{bmatrix} \mathbf{L}_{\boldsymbol{\eta}_t} & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_{\boldsymbol{\phi}_t} \end{bmatrix} \tag{9}$$

provided by the RCD is then introduced in the Kalman filter instead of $\mathbf{C}$. Under the assumption that $\mathbf{A}$, $E[\boldsymbol{\xi}_0]$ and $\boldsymbol{\Sigma}_{\boldsymbol{\xi}_0}$ are known, the first step of the Kalman filter at each time t becomes:

$$\overline{\boldsymbol{\xi}}_{t|t-1} = \mathbf{A}\overline{\boldsymbol{\xi}}_{t-1} + \mathbf{w}_t \tag{10}$$

$$\boldsymbol{\Sigma}_{\overline{\boldsymbol{\xi}}_{t|t-1}} = \mathbf{A}\boldsymbol{\Sigma}_{\overline{\boldsymbol{\xi}}_{t-1}}\mathbf{A}' + \boldsymbol{\Sigma}_{\mathbf{W}}, \tag{11}$$

where $\boldsymbol{\Sigma}_{\mathbf{W}}$ is defined in Section 2.
In the second step we update $\overline{\boldsymbol{\xi}}_{t|t-1}$ on the basis of the new observations $\mathbf{z}_t$:

$$\overline{\boldsymbol{\xi}}_t = \overline{\boldsymbol{\xi}}_{t|t-1} + \overline{\mathbf{K}}_t(\mathbf{z}_t - \overline{\mathbf{z}}_t) \tag{12}$$

$$\boldsymbol{\Sigma}_{\overline{\boldsymbol{\xi}}_t} = (\mathbf{I} - \overline{\mathbf{K}}_t\mathbf{L}_t^{-1})\boldsymbol{\Sigma}_{\overline{\boldsymbol{\xi}}_{t|t-1}}, \tag{13}$$

where $\overline{\mathbf{z}}_t = \mathbf{L}_t^{-1}\overline{\boldsymbol{\xi}}_{t|t-1}$, $\overline{\mathbf{K}}_t = \boldsymbol{\Sigma}_{\overline{\boldsymbol{\xi}}_{t|t-1}}(\mathbf{L}_t')^{-1}\boldsymbol{\Sigma}_{\mathbf{Z}_t}^{-1}$.
The solution $\overline{\boldsymbol{\xi}}_t$ is unique by construction. The indeterminacy of the dynamic solution derives in fact from the arbitrary term $\mathbf{P}_t\boldsymbol{\omega}_t$ in (6), with $\boldsymbol{\Sigma}_{\boldsymbol{\omega}_t} \neq \mathbf{0}$, and the definition of the forecasting $\widehat{\mathbf{z}}_t$ in the Kalman filter by model (2), where $\widehat{\mathbf{z}}_t = \mathbf{C}\widehat{\boldsymbol{\xi}}_{t|t-1}$. In the alternative solution, substituting matrix $\mathbf{C}$ by $\mathbf{L}_t^{-1}$ in the definition of $\widehat{\mathbf{z}}_t$, which we indicate by $\overline{\mathbf{z}}_t$, to distinguish the two cases, allows avoiding the arbitrariness of the "latent" scores, due to both the factorial nature of the measurement models and the dynamics, because the Guttman's result (1955) is no longer appropriate. Guttman considers models like (2), in which the indeterminacy of the $\boldsymbol{\xi}_t$ and $\mathbf{v}_t$ results from the impossibility of identifying m+p basis vectors, where only p observable variables are available. The RCD, on the contrary, provides a definition of the latent variables as a linear combination of the observable variables, so that indeterminacy vanishes. Hence, if $E[\boldsymbol{\xi}_0]$ and $\boldsymbol{\Sigma}_{\boldsymbol{\xi}_0}$ are known, the RCD introduced in the recursive procedure of the Kalman filter obtains a unique approximation for "latent" scores.
Otherwise, since $E[\boldsymbol{\xi}_0]$ and $\boldsymbol{\Sigma}_{\boldsymbol{\xi}_0}$ are often not known, we consider at the first step t=1 the estimate $\overline{\boldsymbol{\xi}}_1$ by (8). Following this procedure we eliminate also the indeterminacy which is due to the not unique estimate of $\boldsymbol{\xi}_1$.
However, the new method proposed cannot be considered a model, because the solutions provided cannot be interpreted as causes of the indicators. From here we derive the necessity of the formulation of a proper Structural Dynamic Model with unique latent scores.

4 A new version of the SDL based on RCFM

In analogy with the proposal of Vittadini and Haagen (2002) for the static case, we present an alternative formulation of the SDL, by extending the RCFM of Haagen and Oberhofer (1999), which we will first describe.
The different assumption of the RCFM to the common FA is that every object i, (i=1,...,N), can be observed R-times (i.e. we have repeated observations for every object). Thus we obtain the following model equation:

$$ {}_{(r)}\mathbf{z}_i = \mathbf{C}\boldsymbol{\xi}_i +_{(r)} \mathbf{v}_i, \qquad r = 1, ..., R; i = 1, ..., N \tag{14}$$

where the r-th repetition is denoted with the index r, ${}_{(r)}\mathbf{z}_i$ is a $p \times 1$ vector of observable variables, $\mathbf{C}$ is a $p \times m$ matrix of factor loadings, $\boldsymbol{\xi}_i$ is a $m \times 1$ vector of common factors and does not depend on r, ${}_{(r)}\mathbf{v}_i$ is a $p \times 1$ vector of specific factors.
Moreover, further assumptions are:

$$p > m, \tag{15}$$
$$rank(\mathbf{C}) = m, \tag{16}$$
$$E[\boldsymbol{\xi}] = \mathbf{0}; E[{}_{(r)}\mathbf{v}] = \mathbf{0}, \qquad r = 1, ..., R \tag{17}$$
$$\boldsymbol{\Sigma}_{\boldsymbol{\xi}} = \mathbf{I}, \tag{18}$$
$$\boldsymbol{\Sigma}_{\boldsymbol{\xi}_{(r)}\mathbf{v}} = \mathbf{0}, \qquad r = 1, ..., R \tag{19}$$
$$\boldsymbol{\Sigma}_{{}_{(r)}\mathbf{v}_{(s)}\mathbf{v}} = \delta_{rs}\mathbf{D}, \mathbf{D} = diag(d_1, ..., d_p), \qquad r, s = 1, ..., R \tag{20}$$

with $d_j > 0, j = 1, ..., p$.
By writing equation (14) in a compact notation, we get the RCFM:

$$ {}_R\mathbf{z} =_R \mathbf{C}\boldsymbol{\xi} +_R \mathbf{v}, \tag{21}$$

where ${}_R\mathbf{z} = ({}_{(1)}\mathbf{z}', ...,{}_{(R)} \mathbf{z}')'$ $(pR \times 1)$, ${}_R\mathbf{C} = (\mathbf{C}', ..., \mathbf{C}')'$ $(pR \times m)$, $\boldsymbol{\xi} = (\xi_1, ..., \xi_m)'$ $(m \times 1)$, ${}_R\mathbf{v} = ({}_{(1)}\mathbf{v}', ...,{}_{(R)} \mathbf{v}')'$ $(pR \times 1)$.
Equation (21) represents a Common Factor Model with pR observable variables and m common factors; the number of parameters is fixed.
Haagen and Oberhofer (1999) demonstrated that, for given $\mathbf{C}$ and $\mathbf{D}$, the indeterminacy vanishes as $R \to \infty$, so that RCFM solves the indeterminacy of common factors asymptotically. In fact Haagen and Oberhofer (1999) demonstrated that in the RCFM

$$ {}_R\widehat{\boldsymbol{\xi}} =_R \mathbf{C}'\boldsymbol{\Sigma}^{-1}_{{}_R\mathbf{Z}} \; {}_R\mathbf{z}, \tag{22}$$

e.g. the regression estimator of factor scores, can always be written as:

$$ {}_R\widehat{\boldsymbol{\xi}} = \mathbf{C}'(\mathbf{C}\mathbf{C}' + \frac{1}{R}\mathbf{D})^{-1}\overline{\mathbf{z}}, \tag{23}$$

with $\overline{\mathbf{z}} = \frac{1}{R}\sum_{r=1}^{R} {}_{(r)}\mathbf{z}$, and converges to $\boldsymbol{\xi}$ in quadratic mean as $R \to \infty$.
Moreover they demonstrated that the covariance matrix of the arbitrary part

$\boldsymbol{\omega}_t$, i.e. the arbitrariness $\boldsymbol{\Sigma}_{\boldsymbol{\omega}_t} \to 0$ as $R \to \infty$.
Assuming that at every time interval t, (t=1,...T), each object i, (i=1...,N), can be observed R-times on vector $\mathbf{z}_t$, we reformulate the SDL through the RCFM as follows (with no distributional assumptions):

$$\boldsymbol{\xi}_t = \mathbf{A}\boldsymbol{\xi}_{t-1} + \mathbf{w}_t \qquad t = 1, ..., T \tag{24}$$
$$_R\mathbf{z}_t =_R \mathbf{C}\boldsymbol{\xi}_t +_R \mathbf{v}_t, \qquad t = 1, ..., T \tag{25}$$

with $\boldsymbol{\xi}_t = [\boldsymbol{\eta}_t, \boldsymbol{\phi}_t]$ $(m \times 1)$, $\mathbf{A}$ $(m \times m)$, $\mathbf{w}_t$ $(m \times 1)$, $_R\mathbf{z}_t = [_{(1)}\mathbf{z}'_t, ...,_{(R)} \mathbf{z}'_t]'$ $(pR \times 1)$, $_R\mathbf{C} = (\mathbf{C}', ..., \mathbf{C}')'$ $(pR \times m)$, $_R\mathbf{v}_t = (_{(1)}\mathbf{v}'_t, ...,_{(R)} \mathbf{v}'_t)'$ $(pR \times 1)$, $\boldsymbol{\xi}_t = (\xi_{t1}, ..., \xi_{tm})'$ $(m \times 1)$.
At time t=1 we propose estimating $\boldsymbol{\xi}_1$ by means of (22). At time t, (t=2,...,T), supposing that $\mathbf{A}$ and $_R\mathbf{C}$ are known, $\boldsymbol{\xi}_t$ is estimated by means of the Kalman filter, which in the first step becomes:

$$\widehat{\boldsymbol{\xi}}_{t|t-1} = \mathbf{A}\widehat{\boldsymbol{\xi}}_{t-1} + \mathbf{w}_t \tag{26}$$
$$\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\xi}}_{t|t-1}} = \mathbf{A}\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\xi}}_{t-1}}\mathbf{A}' + \boldsymbol{\Sigma}_{\mathbf{W}}. \tag{27}$$

In the second step $\widehat{\boldsymbol{\xi}}_{t|t-1}$ is updated on the basis of the new observations $_R\mathbf{z}_t$:

$$\widehat{\boldsymbol{\xi}}_t = \widehat{\boldsymbol{\xi}}_{t|t-1} +_R \mathbf{K}_t(_R\mathbf{z}_t -_R \widehat{\mathbf{z}}_t) \tag{28}$$
$$\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\xi}}_t} = (\mathbf{I} - {_R\mathbf{K}_t}{_R\mathbf{C}})\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\xi}}_{t|t-1}}, \tag{29}$$

where $_R\widehat{\mathbf{z}}_t =_R \mathbf{C}\widehat{\boldsymbol{\xi}}_{t|t-1}$ and $_R\mathbf{K}_t = \boldsymbol{\Sigma}_{\widehat{\boldsymbol{\xi}}_{t|t-1}}{_R\mathbf{C}'}\boldsymbol{\Sigma}^{-1}_{_R\mathbf{Z}_t}$ of dimension $(m \times pR)$.
As $R \to \infty$ the solution $\widehat{\boldsymbol{\xi}}_1$ is unique and satisfies the fundamental hypothesis of the FA indicated in (19), as we demonstrate in the following.
In fact, with reference to model (21) we have:

$$_R\mathbf{C} =_R \mathbf{z}_1\boldsymbol{\xi}'_1(\boldsymbol{\xi}'_1\boldsymbol{\xi}_1)^{-1} =_R \mathbf{z}_1\boldsymbol{\xi}'_1. \tag{30}$$

Then, under the hypothesis that $R \to \infty$ and substituting $\boldsymbol{\xi}_1$ by the (22) and $_R\mathbf{C}$ by the (30), equation (21) can be rewritten as:

$$\begin{aligned}
_R\mathbf{z}_1 &=_R \mathbf{C}\boldsymbol{\xi}_1 +_R \mathbf{v}_1 = \\
&=_R \mathbf{C}{_R\mathbf{C}'}\boldsymbol{\Sigma}^{-1}_{_R\mathbf{Z}_1}{_R\mathbf{z}_1} + (_R\mathbf{z}_1 -_R \mathbf{C}{_R\mathbf{C}'}\boldsymbol{\Sigma}^{-1}_{_R\mathbf{Z}_1}{_R\mathbf{z}_1}) = \\
&=_R \mathbf{z}_1\boldsymbol{\xi}'_1\boldsymbol{\xi}_1{_R\mathbf{z}'_1}\boldsymbol{\Sigma}^{-1}_{_R\mathbf{Z}_1}{_R\mathbf{z}_1} + (_R\mathbf{z}_1 -_R \mathbf{z}_1\boldsymbol{\xi}'_1\boldsymbol{\xi}_1{_R\mathbf{z}'_1}\boldsymbol{\Sigma}^{-1}_{_R\mathbf{Z}_1}{_R\mathbf{z}_1}) = \\
&=_R \mathbf{z}_1\mathbf{P}_{\boldsymbol{\xi}_1}\mathbf{P}_{\mathbf{Z}_1} + (_R\mathbf{z}_1 -_R \mathbf{z}_1\mathbf{P}_{\boldsymbol{\xi}_1}\mathbf{P}_{\mathbf{Z}_1}) = \\
&=_R \mathbf{z}_1\mathbf{P}_{\boldsymbol{\xi}_1} + (_R\mathbf{z}_1 -_R \mathbf{z}_1\mathbf{P}_{\boldsymbol{\xi}_1}) = \\
&=_R \mathbf{z}_1\mathbf{P}_{\boldsymbol{\xi}_1} +_R \mathbf{z}_1\mathbf{Q}_{\boldsymbol{\xi}_1},
\end{aligned} \tag{31}$$

where $\mathbf{P}_{\boldsymbol{\xi}_1}$ is the projector onto the space spanned by $\boldsymbol{\xi}_1$ and $\mathbf{Q}_{\boldsymbol{\xi}_1} = \mathbf{I} - \mathbf{P}_{\boldsymbol{\xi}_1}$.
Consequently, as $R \to \infty$, both causes of indeterminacy of the SDL, indicated in Section 3, vanish and the estimates for the latent variables, provided by

the Kalman filter, become unique.
By the end, it should be noted that the (15) and the (16) are also fundamental hypotheses of the FA, while the replicability of the observations and the (20) are the basis for the RCFM. The (17) and the (18) are instead not essential; the model proposed can be surely extended to a more general case.
We conclude that, by using the new formulation of the SDL expressed in the (22)-(23) and under the assumption to consider at the first step $t = 1$ the estimator $\widehat{\boldsymbol{\xi}}_1$ defined by (22), the indeterminacy of the SDL is definitely overcome.

5 The Application

The application regards an experiment of the Department of Physics, University of Milan. The goal of the experiment is the measurement of the system temperature of a radiometer for microwaves astronomy. Different level of system temperature are observed at different times due to the effect of different operating temperature. At each time t the experiment is repeated 2,000 times at the same conditions, i.e. the observed variable, which measures the "true" variable with white noise, is collected 2,000 times.
In order to obtain, at each time t, unique values for the "true" measure underlying the observations, the model proposed in (24) and (25) is applied. The application corresponds to the theoretical issues of the model proposed. In fact, first of all we have, at the same time t, several replications (under equal conditions) of the same observed variable. Therefore, the assumptions inherent to the replicability of observations are respected. Secondly, the measurement model (25) is a model with errors in variables, (i.e. a particular case of the FA), with replicated observations. By the end, the (24) represents the relation between the measure of interest at time t and the same measure at time t-1. For sake of simplicity, we have supposed that passing from time t to t-1 occurs with a constant change of temperature, i.e. the "true" measure at time t differs from the "true" measure at time t-1 of a constant, which represents the change of temperature between time t and t-1.

6 Conclusions

The model proposed, which provides unique latent scores in a dynamic context, seems to be a reasonable alternative to the SDL, because it not only represents a statistical model, but it definitely overcomes the latent score indeterminacy.
The question is whether the consideration $R \to \infty$ is realistic. The main interesting issue is that the result of Haagen and Oberhofer (1999) is not only valid for $R \to \infty$, but also for finite R, if R increases. Simulation studies to verify the empirical validity of the RCFM show that the estimates converge

for R=50 (Taufer, 1992).
The applicability of this model to real problems is surely limited by the assumption that the vectors of latent variables do not depend on replications. A field of application always compatible with the assumption is the case of physical experiments, where it is not difficult to produce a large number of replications.

References

BORDIGNON, S. and TRIVELLATO, U. (1992): Errori nelle variabili e variabili latenti in modelli strutturali stocastici: un quadro di riferimento. *Statistica, LII, 3, 325–345.*

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

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

HAAGEN, K. and OBERHOFER, W. (1999): The replicated common factor-analysis model. *Metron, LVII, 3-4, 35–45.*

HAAGEN, K. and VITTADINI, G. (1991): Regression component decomposition in structural analysis. *Communications in Statistics - Theory and Method, 20, 4, 1153–1161.*

HAAGEN, K. and VITTADINI, G. (1994): Sul problema dell'osservabilità nei modelli strutturali dinamici. In: *Atti XXXVII Riun. Scient. SIS (Sessioni Specializzate).* Sanremo, 247–258.

MINOTTI, S.C. (2002): Some Structural Dynamic Models Overcoming Latent Score Indeterminacy. In: *Atti XLI Riun. Scient. SIS (Sessioni Spontanee).* Cleup, Padova, 423–426.

OTTER, P.W. (1985): Dynamic Feature Space Modelling, Filtering and Self-Tuning Control of Stochastic Systems: A Systems Approach with Economic and Social Applications. In: M. Beckmann and W. Krelle (Eds.): *Lecture Notes in Economics and Mathematical Systems, Vol.246.* Springer, Berlin.

OTTER, P.W. (1992): Dynamic models with latent variables from a system theoretic perspective: theory and applications. *Statistica, LII, 3, 347–363.*

SCHOENEMANN, P.H. and STEIGER, J. (1976): Regression Component Analysis. *British Journal of Mathematical and Statistical Psychology, 29, 175–189.*

SCHOENEMANN, P.H. and HAAGEN, K. (1987): On the use of Factor Scores for Prediction. *Biometrical Journal, 29, 7, 835–847.*

TAUFER, E. (1992): Studio sulla validità empirica di un modello di analisi fattoriale con osservazioni ripetute. *Quaderni di Statistica e Matematica applicata alle Scienze Economico-Sociali, XIV, 1-2, 127–140.*

VITTADINI, G. (1989): Indeterminacy Problems in the Lisrel Model. *Multivariate Behavioral Research, 24, 4, 397–414.*

VITTADINI, G. (1999): Analysis of Qualitative Variables in Structural Models with Unique Solutions. In: M. Vichi and O. Opitz (Eds.): *Classification and Data Analysis: Theory and Application.* Springer, Berlin, 203–210.

VITTADINI, G. and HAAGEN, K. (2002): A Causal Model with Latent Variables and Unique Solutions. In: *Atti XLI Riun. Scient. SIS (Sessioni Specializzate).* Cleup, Padova, 211–220.