UNIVERSITÄT POTSDAM
Wirtschafts- und Sozialwissenschaftliche Fakultät

STATISTISCHE DISKUSSIONSBEITRÄGE

Nr. 1

Hans Gerhard Strohe

Dynamic Latent Variables Path Models

- An Alternative PLS Estimation -

Potsdam 1995
ISSN 0949-068X

STATISTISCHE DISKUSSIONSBEITRÄGE

Nr. 1

Hans Gerhard Strohe

Dynamic Latent Variables Path Models

- An Alternative PLS Estimation -

Herausgeber: Lehrstuhl Statistik (Prof. Dr. Hans Gerhard Strohe)
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Universität Potsdam
Postfach 90 03 27
D-14439 Potsdam
Tel. (+49 331) 977-32 25
Fax. (+49 331) 977-32 10
1995
ISSN 0949-068X

Hans Gerhard Strohe

Dynamic Latent Variables Path Models
- An Alternative PLS Estimation -

Abstract
In this paper a partial least squares (PLS) approach to dynamic modelling with latent variables is proposed. Let **Y** be a matrix of manifest variables and **H** the matrix of the corresponding latent variables. And let **H** = **BH**+ε be a structural PLS model with a coefficient matrix **B**. Then this model can be made a dynamic one by substituting for **B** a matrix **F** = **B** + **C**L containing the lag operator L. Then the structural dynamic model **H** = **FH**+ε is formally estimated like an ordinary PLS model. In an exploratory way the model can be used for forecasting purposes. The procedure is being programmed in ISP.
Key words: PLS, dynamic models, path models

1. Introduction

Linear models with latent variables have become quite common in psychometrics, sociometrics and econometrics. Beside the favourite estimation and confirmation technique LISREL by Jöreskog and Sörbom (e g 1987), the partial least squares PLS algorithm by H. Wold (1973) has gained popularity during the last years despite a certain weakness of theoretical foundations. Particularly, chemometrics have been bringing many contributions to the developement of new methods and new applications of PLS regression (e g Marengo 1991, Hellan et al 1991).
Path modelling with latent variables has several different historical sources. Spearman (1927) introduced group factors into statistical models and Thurstone (1935) generalized Spearmans factor model to more than one common factor. H. Hotelling (1933) and T. Kelley (1935) rediscovered the principal component method. Again H. Hotelling (1935) was the originator of what is called now "canonical correlation" and what he had called "the most predictable criterion".
Sewall Wright (1934) contributed probably the term "path analysis" but without taking into account latent variables. Nevertheless, his work had considerable impact on econometric and biometric modelling. The term "latent" in connection with variables to be modelled was probably first used by Paul Lazarsfeld in 1950. Jöreskog (1970) found the LISREL approach (Jöreskog /Sörbom 1987) to linear path models with latent variables. LISREL is model oriented and rather confirmative.
The PLS approach to path models has been introduced by H. Wold (e g 1973), the model being defined purely by an algorithm. It is data oriented and rather descripitve or explorative. Only recently its mathematical and statistical properties have become more apparent particulary in one and two-block models (Helland 1988, Lohmöller 1989, Schneeweiß 1993).
Stone and Brooks (1989) have shown "that, with a particular objective criterion for the

construction, the procedures of ordinary least squares and principal component regression occupy the oposite ends of a continous spectrum, with partial least squares lying in between". We do not very often find latent variable path models in the domain of time series. But there are some remarkable exceptions: Apel and Lohmöller (1992) intruduced time in their ecology model as a particular latent variable. Regarding the treatment of the other variables, this PLS time series model did not differ from PLS cross section models.
A specific time series approach to path modelling can be found with Hillmer (1993). Here the manifest variables are submitted to an ARIMA analysis before the residuals or innovations are composed to latent variables of LISREL models. The aim of these models was testing causal relationships rather than forecasting underlying processes.
Otter (1992) presented a very stimulating approach to time series that inherently combined dynamic modelling with latent variable path analysis. He formulated a generalization of Jöreskog's structural LISREL model to a dynamic one. He stated that the LISREL model and the linear stochastic state space model basically have the same structure. He gave some examples in which the method of state space models and Kalman filter were used for estimating dynamic models with latent variables.

In section 2 of this paper a short description of the main steps of H. Wold's PLS estimation algorithm for latent variable path models is given.
Then, in section 3, a PLS-like approach to a certain broad class of dynamic models with latent variables will be proposed. The way for predicting by these models and a measure for goodness of fit is deduced.
Finally a small two-block model with an autoregressive distributed lag relation between the latent variables is presented.

2. Path Models and the Classical Partial Least Squares Algorithm

A path model (e g Rönz/Strohe 1994) involves M manifest variables y^m (m=1, ..., M) and K<M latent variables η^k (k=1,...,K). The latent variables are not directly observable. They are assumed to be certain constructs built from observable manifest variables. Furthermore, the latent variables are assumed to be connected by linear relations.The system of these relationships is called the inner model:

$$\boldsymbol{\eta}_t = \mathbf{b}_0 + \mathbf{B}\boldsymbol{\eta}_t + \boldsymbol{\nu}_t \tag{1}$$

where $\boldsymbol{\eta}_t = (\eta_t^1, \eta_t^2, ..., \eta_t^K)'$ is the column vector of the scores of all latent variables $\eta^1, ..., \eta^K$ for one certain case or time t (t = 1, ..., T).
B is a triangular matrix of path coefficients with zero diagonal and $\mathbf{b}_0$ is a location parameter vector usually set equal zero. The error term $\boldsymbol{\nu}_t$ has zero expectation and the covariance matrix cov($\boldsymbol{\nu}$)=$\boldsymbol{\Psi}$.
The outer or measurement model describes the assumed linear relations between the observable manifest and the inobservable latent variables. Firstly we have the loading relation:

$$\mathbf{y}_t = \mathbf{p}_0 + \mathbf{P}\boldsymbol{\eta}_t + \boldsymbol{\varepsilon}_t \tag{2}$$

with a matrix of path coefficients $\mathbf{P}$ and a zero expectation disturbance term $\boldsymbol{\varepsilon}_t$, $\mathrm{cov}(\boldsymbol{\varepsilon}_t)=\boldsymbol{\Theta}$. Again the location parameter $\mathbf{p}_0$ is usually set to zero because the manifest variables are regarded as differences from their means.
In the case of a partial least squares path model, secondly the weight relation

$$\boldsymbol{\eta}_t = \mathbf{W}'\,\mathbf{y}_t + \boldsymbol{\delta}_t \tag{3}$$

belongs to the outer model, where $\mathbf{W}$ is a block diagonal weight matrix, e g

$$W = \begin{pmatrix} \omega_{11} & 0 & \dots & \dots & 0 \\ \omega_{12} & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \omega_{1m_1} & 0 & \dots & \dots & 0 \\ 0 & \omega_{21} & \dots & \dots & 0 \\ 0 & \omega_{22} & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \omega_{2m_2} & \dots & \dots & 0 \\ \dots & \dots & \dots & 0 & \omega_{k1} \\ \dots & \dots & \dots & 0 & \omega_{k2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & \dots & \dots & 0 & \omega_{km_k} \end{pmatrix}$$

and $\boldsymbol{\delta}_t$ an zero mean error term. In Wold's basic model $\boldsymbol{\delta}_t$ is assumed to be constantly zero. We follow this assumption. That means, the latent variables are taken to be undisturbed weighted sums of manifest variables.

The iterative estimation of the weights $\mathbf{W}$ is the main aim of Wold's PLS algorithm.
In order to be brief in the following description we will use the same symbols used above for the "true" variables and parameters for the stepwise estimations of these variables and parameters.

The procedure is to start with more or less arbitrarily chosen weights $\mathbf{W}$. From this the starting scores of the latent variables are obtained:

$$\boldsymbol{\eta}_t := \mathbf{W}'\,\mathbf{y}_t \tag{4}$$

the data vectors $\mathbf{y}_t$ ($t = 1, \ldots, T$) being given. Then each of these variables are standardized to unit variance over all cases.

$$\eta_t^k := \frac{\eta_t^k}{\sqrt{var(\eta^k)}}$$

where var($\boldsymbol{\eta}^k$) is the empirical variance of the scores $\eta_1^k, ..., \eta_T^k$.
Then follows the calculation of an inner approximation $\boldsymbol{\eta}^{k*}$ for the latent variables $\boldsymbol{\eta}^k$ by the variables in their "neighbourhood" according to the paths in the inner model, i.e. $\boldsymbol{\eta}^{k*}$ is a weighted sum of all latent variables on which $\boldsymbol{\eta}^k$ either depends or which depends on $\boldsymbol{\eta}^k$.
Finally new weights $\mathbf{W} = (\omega_{mk})$ are estimated, e g by OLS regression out of the relation $y_t^m = \omega_{mk}\, \eta_t^{k*} + \nu_t^{mk}$ if $\mathbf{y}^m$ is a manifest variable belonging to the block of the latent variable $\boldsymbol{\eta}^k$ (mode A of H. Wold).
New scores of the latent variables arise out of these weights and the iteration cycle starts again with (4). The iteration is stopped when the latent variable scores do not change relevantly according to a given criterion.
Now the scores of the latent variables being known, the parameter matrices **B** and **P** are easily estimated by OLS, considering the inner and outer model equation (1) and (2) respectively as regression equations.

The residuals of the model estimated this way can be treated the same way as the original manifest variables in order to obtain a second dimension of latent variables similar to the second principal component in factor analysis.
More details about PLS are given by Lohmöller (1989), who is the author of the PLS computer program LVPLS (Lohmöller 1984).

3. A Dynamic PLS Model (DPLS)

3.1. The structure

The inner PLS model can be written in a structural form

$$\boldsymbol{\eta}_t = \mathbf{B}\boldsymbol{\eta}_t + \boldsymbol{\nu}_t \qquad (1^*)$$

where $\boldsymbol{\eta}_t$ is a vector of K latent variables with zero mean and unit variance and $\boldsymbol{\nu}_t$ a disturbance term with zero expectation and covariance $\boldsymbol{\Psi}$. The index t = 1, ..., T now denotes time. The K×K-matrix **B** is considered to be triangular i.e. the model is to be recursive.
It is further assumed that η_t can be measured by the observable variable vector y_t (outer PLS model or measurement model)

$$\mathbf{y}_t = \mathbf{P}\boldsymbol{\eta}_t + \boldsymbol{\varepsilon}_t, \qquad (2^*)$$

where **P** is a M×K matrix of loadings. By ordering the elements of $\mathbf{y}_t$ in the way that the first

m_1 elements belong to $\boldsymbol{\eta}_t^1$, the next m_2 to $\boldsymbol{\eta}_t^2$ and so on, the matrix **P** can be made block diagonal i.e. the first m_1 lines of **P** contain values different from 0 only in the first column, the next m_2 lines contain elements $\neq 0$ in the second column and so on. The error term $\boldsymbol{\varepsilon}_t$ is white noise with zero mean and covariance $\boldsymbol{\Theta}$.

The structural model (1*) can be made dynamic by implementing a lagged term:

$$\boldsymbol{\eta}_t = \mathbf{B}\boldsymbol{\eta}_t + \mathbf{C}\boldsymbol{\eta}_{t-1} + \boldsymbol{\nu}_t \tag{5}$$

where **C** is a K×K matrix of coefficients and $E(\boldsymbol{\eta}_0) = \boldsymbol{\mu}$, $\text{cov}(\boldsymbol{\eta}_0) = \boldsymbol{\Sigma}_{\eta^0}$ are additional start conditions for the data generating process.
Equation (5) can be rewritten as

$$(\mathbf{I}\text{-}\mathbf{B})\,\boldsymbol{\eta}_t = \mathbf{C}\boldsymbol{\eta}_{t-1} + \boldsymbol{\nu}_t$$

or in order to obtain the reduced form

$$\boldsymbol{\eta}_t = (\mathbf{I}\text{-}\mathbf{B})^{-1}\mathbf{C}\boldsymbol{\eta}_{t-1} + (\mathbf{I}\text{-}\mathbf{B})^{-1}\,\boldsymbol{\nu}_t$$

or

$$\boldsymbol{\eta}_t = \mathbf{D}\boldsymbol{\eta}_{t-1} + \mathbf{u}_t \tag{6a}$$

$$\mathbf{y}_t = \mathbf{P}\boldsymbol{\eta}_t + \boldsymbol{\varepsilon}_t \tag{6b}$$

where **D** is defined as $(\mathbf{I}\text{-}\mathbf{B})^{-1}\mathbf{C}$ and $\mathbf{u}_t = (\mathbf{I}\text{-}\mathbf{B})^{-1}\boldsymbol{\nu}_t$ is an error term with $E(\mathbf{u}_t) = 0$ and

$$\text{cov}(\mathbf{u}_t) = (\mathbf{I}\text{-}\mathbf{B})^{-1}\,\boldsymbol{\Psi}\,((\mathbf{I}\text{-}\mathbf{B})^{-1})'. \tag{7}$$

The transformed model (6a,b) corresponds formally to state space models (Aoki 1990) with state vector $\boldsymbol{\eta}_t$, an observable time series vector $\mathbf{y}_t$ and noise vectors $\mathbf{u}_t$ and $\boldsymbol{\varepsilon}_t$. Hence Otter (1992) proposed state space specific methods, particularly the Kalman filter, for its treatment, which gives satisfactory results under rather restrictive distribution conditions. A similar approach called DYMIMIC was presented by Bordignon / Gaetan (1989) and Trivellato et al (1993).

Deviating from Otters approach which has some analogy with Jöreskogs LISREL we try and generalize here Wold's PLS. In order to obtain a PLS path model, the weight relation

$$\boldsymbol{\eta}_t = \mathbf{W}'\,\mathbf{y}_t \tag{8}$$

has to be added to the model (6).

Returning to equation (5), the dynamic form of the structural model can be transformed into the shape of the "normal" PLS model:

$$\boldsymbol{\eta}_t = \mathbf{F}\boldsymbol{\eta}_t + \boldsymbol{\nu}_t \tag{9}$$

where

$$\mathbf{F} = \mathbf{B} + \mathbf{C}L \tag{10}$$

is a matrix containing the lag operator L defined by $Lx_t = x_{t-1}$.
On what we call now the dynamic PLS model (DPLS)

$$\boldsymbol{\eta}_t = \mathbf{F}\boldsymbol{\eta}_t + \mathbf{v}_t, \tag{11}$$
$$\mathbf{y}_t = \mathbf{P}\boldsymbol{\eta}_t + \boldsymbol{\varepsilon}_t \tag{12}$$

the classical PLS alogrithm is formally applicable. In advance, Boolean design matrices $\mathbf{D}_B$, $\mathbf{D}_C$ and $\mathbf{D}_P$ corresponding to the unlagged and lagged dependencies in the inner model (11) and to the outer model (12), i e to the restrictions for the coefficient matrices **B, C** and **P** , are to be fixed. The inner model (11) can be illustrated by a path diagramme additionally including arrows for the lagged relationships (dotted arrows in fig. 2). The inner design matrix $\mathbf{D}_B$ contains unities where there is a connection between two latent variables in the path model and consists of zeros elsewhere. In analogy to this, the lag design matrix $\mathbf{D}_C$ consists of unities and zeros corresponding to whether or not there is first order lagged (auto-)regression between latent variables assumed to exist. $\mathbf{D}_P=[d_{mk}]$ is the outer design matrix corresponding to whether or not a variable $\mathbf{y}^m$ of **Y** belongs to the block of a certain latent variable i.e. a row $\boldsymbol{\eta}^k$ of **H**, e g

$$D_P = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

3.2. Partial Least Squares Estimation with Dynamic Inner Approximation

In this section the symbols for the empirically estimated latent variables and coefficients are not distinguished from those for the corresponding theoretical quantities. There is not any

cause for confusion as the latter do not appear in this section.
In order to estimate the weight matrix **W** the following steps are to be executed:

i. First representation of the latent variables as components of the manifest variables with chosen starting values for the matrix **W**

$$\boldsymbol{\eta}_t = \mathbf{W}'\mathbf{y}_t \tag{13}$$

ii. Standardization of the latent variables to unit variance

$$\boldsymbol{\eta}_t := \sqrt{T}(\mathbf{I} * \mathbf{HH}')^{-1/2}\boldsymbol{\eta}_t \tag{14}$$

where **H** is the K×T matrix of all time scores of the vector $\boldsymbol{\eta}_t$ for t = 1, ..., T. Thus the diagonal elements of **HH**'/T are the empirical variances of the individual variables in the vector $\boldsymbol{\eta}_t$. Element-wise multiplication by a unit matrix **I** is denoted by *.

iii. Calculation of what is called "environment" variables corresponding to the inner path model and taking into account each relationship for both the dependent and the independent variables:

$$\begin{aligned} \boldsymbol{\eta}_t^* &= \mathbf{F}^*\boldsymbol{\eta}_t \\ &= (\mathbf{B}^* + \mathbf{C}^*L + \mathbf{C}^{0\prime}L^{-1})\boldsymbol{\eta}_t \\ &= \mathbf{B}^*\boldsymbol{\eta}_t + \mathbf{C}^*\boldsymbol{\eta}_{t-1} + \mathbf{C}^{0\prime}\boldsymbol{\eta}_{t+1} \end{aligned} \tag{15}$$

where $\mathbf{B}^*$ and $\mathbf{C}^*$ are suitable inner weighting matrices, e g

$$\mathbf{B}^* = (\mathbf{D}_B + \mathbf{D}_B')*\mathbf{R} \tag{16}$$
$$\mathbf{C}^* = \mathbf{D}_C*\mathbf{A} \tag{17}$$
$$\mathbf{C}^0 = \mathbf{D}_C^0 *\mathbf{A} \tag{18}$$

with $\mathbf{D}_B$ and $\mathbf{D}_C$ being the design matrices for the inner model and $\mathbf{D}_C^0$ denoting the design matrix $\mathbf{D}_C$ with diagonal elements set equal zero. The matrices

$$\mathbf{R} = \mathbf{HH}'/T \tag{19}$$
$$\mathbf{A} = \mathbf{H}(L\mathbf{H})'/T \tag{20}$$

are the correlation matrix and an approximation for the first order autocorrelation matrix of the latent variables, respectively, with

$$\begin{aligned} \mathbf{H} &= (\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, ..., \boldsymbol{\eta}_T) \quad \text{and hence} \\ L\mathbf{H} &= (\boldsymbol{\eta}_0, \boldsymbol{\eta}_1, ..., \boldsymbol{\eta}_{T-1}). \end{aligned} \tag{21}$$

The starting condition $\boldsymbol{\eta}_0$ is to be drawn from a population with $E(\boldsymbol{\eta}_0)=\mathbf{0}$ and $var(\boldsymbol{\eta}_0)=1$.

iv. A new estimation of the weight matrix **W** is gained by OLS estimating the parameters

of the regression equation
$y_t^m = \omega_{mk}\, \eta_t^{k*} + \nu_t^{mk}$ if d_{mk}=1 (22)
where $\mathbf{D}_P$ =[d_{mk}] is the outer design matrix corresponding to whether or not a variable $\mathbf{y}^m$ of **Y** is connected to a row $\boldsymbol{\eta}^k$ of **H**.
From this follows in analogy to H. Wold's mode B the OLS estimator

$$\omega_{mk} = \frac{\eta^{k*}\, y^{m\prime}}{\eta^{k*}\, \eta^{k*\prime}} \tag{23}$$

v. These regression coefficients ω_{mk} are substituted into the weight matrix **W**
W: = (ω_{mk}).

Using this new weight matrix the procedure continues by repeating step i. The iteration process is considered to converge when subsequent estimations of the latent variables scores $\boldsymbol{\eta}_t$ in step ii. do not relevantly differ from the previous ones.
Then the coefficient matrices **B** and **C** of the inner model (11)

$$\boldsymbol{\eta}_t = (\mathbf{B} + \mathbf{C}\mathrm{L})\, \boldsymbol{\eta}_t + \boldsymbol{\nu}_t$$

are to be estimated by a suitable estimation method for dynamic models depending on the dynamic properties of the latent variables (e g Banerjee ... Hendry 1993). The search for a method should include diagnostic tests of the residuals as well as cross validation of the predictive relevance.
The loadings **P** of the outer model (12)

$$\mathbf{y}_t = \mathbf{P}\, \boldsymbol{\eta}_t + \boldsymbol{\varepsilon}_t$$

are estimated by simple OLS:
$\mathbf{P} = \mathbf{YH}'\, (\mathbf{HH}')^{-1}$.

The dynamic partial least squares algorithm here described is being programmed in the command language of the interactive statistic processor PC-ISP/DGS (© Datavision AG, Schweiz).

3.3. Prediction and Redundancy

Here again, the estimated coefficient matrices **H**, **B**, **C**, **P** and **W** are taken for the "true" ones. By substituting (11) for $\mathbf{y}_t$ in (12) we obtain

$$\mathbf{y}_t = \mathbf{P}\boldsymbol{\eta}_t + \boldsymbol{\varepsilon}_t$$

$$= \quad \mathbf{PF\eta_t + P\nu_t + \varepsilon_t} \tag{24}$$

Then substituting (8) for $\boldsymbol{\eta}_t$, we have

$$\begin{aligned} \mathbf{y}_t &= \mathbf{PFW'y_t + P\nu_t + \varepsilon_t} \\ &= \mathbf{P(B{+}CL)W'y_t + P\nu_t + \varepsilon_t} \end{aligned} \tag{25}$$

Hence, the identity (10) gives the prediction formula

$$\mathbf{y}_t = [\mathbf{PBW'y_t + P\,C\,W'\,y_{t-1}}] + [\mathbf{P\nu_t + \varepsilon_t}] \tag{26}$$

with $\mathbf{W'y_0} \equiv \boldsymbol{\eta}_0$ chosen as mentioned in sc. 3.2 step iii.

Now we consider the construction of a goodness-of-fit criterion that provides a validation measure for the predictive relevance of the model. From (24) follows that the predictable part of $\mathbf{y}_t$ is $\mathbf{y}_t^* = \mathbf{P\,F\,\eta_t}$. Let $\mathbf{Y}^* = (\mathbf{y}_1^*, ..., \mathbf{y}_T^*)$ denote the whole predicted data matrix. Then the covariance of these predictions is

$$\begin{aligned} \mathbf{cov(Y^*)} &= \mathbf{P\,F\,cov(H)\,B'\,P'} \\ &= \mathbf{P\,F\,R\,F'\,P'} \end{aligned} \tag{27}$$

with $\mathbf{R}$ being the correlation or covariance matrix of the latent variables $\mathbf{H} = (\boldsymbol{\eta}_1, ..., \boldsymbol{\eta}_T)$ and

$$\begin{aligned} \mathrm{T}\mathbf{cov(Y^*)} &= \mathbf{P(B{+}CL)\,H\,H'\,(B{+}CL)'\,P'} \\ &= \mathbf{P(B{+}CL)\,H\,(P(B{+}CL)\,H')'} \\ &= \mathbf{(PB{+}PCL)\,H\,[(PB{+}PCL)\,H]'} \\ &= \mathbf{PBHH'B'P' + PC(LH)H'B'P' + PBH(LH)'C'P' + PC(LH)(LH)'C'P'} \\ &\approx 2\,\mathrm{T}\,\mathbf{(PBRB'P' + PCAC'P')} \end{aligned}$$

with $\mathbf{A}$ being the first order autocorrelation matrix. The slight fuzziness of this relation arises from that tiny differences might occur between the covariances of the latent variables $\mathbf{HH'}$ / T and those of the lagged latent variables $\mathbf{(LH)(LH)'}$/ T. Using the notation

$$\mathbf{cov(Y^*)} \approx 2\,\mathbf{(PBRB'P' + PCAC'P') = G^*} \tag{28}$$

it is easy to see that $\mathbf{G}^*$ contains in its diagonal the variances of the predictable or what Lohmöller (1989) calls the redundant part of the manifest variables.
Following further Lohmöller we calculate the ratio of two diagonal matrices

$$\mathbf{G} = \mathbf{(I * G^*)\,(I * \Sigma_y)^{-1}} \tag{29}$$

where $\mathbf{\Sigma}_y$ denotes the empirical covariance matrix of the manifest variables. The figures in the diagonale of G are proportions expressing to which extent the variance of each manifest variable is reproduced by the variance of the predictable part, i. e. by the model.

The average of these measures

$$G^2 \quad = \quad \text{trace}\ \mathbf{G} / M \tag{30}$$

is the redundancy coefficient or average redundancy and is used for the evaluation of the goodness of fit of the model as a whole.

3.4. An example

For simplicity an example of a two latent variables model is considered:

$$\begin{aligned} \boldsymbol{\eta}_t^{2} &= b\ \boldsymbol{\eta}_t^{1} + c_1\boldsymbol{\eta}_{t-1}^{1} + c_2\boldsymbol{\eta}_{t-1}^{2} + \boldsymbol{\nu}_t && \\ y_t^{m} &= p_{m1}\ \boldsymbol{\eta}_t^{1} + \boldsymbol{\varepsilon}_t^{m} && \text{for} \quad m = 1,\dots,4 \\ y_t^{m} &= p_{m2}\ \boldsymbol{\eta}_t^{2} + \boldsymbol{\varepsilon}_t^{m} && \text{for} \quad m = 5,\dots,7 \end{aligned} \tag{31}$$

Without the lagged terms this would be a simple case of PLS regression recently discussed and analysed very often in chemometric literature. But under consideration of its dynamic structure the model is a bit more complex.

A suitable path diagramm for this model could be that in fig. 1.

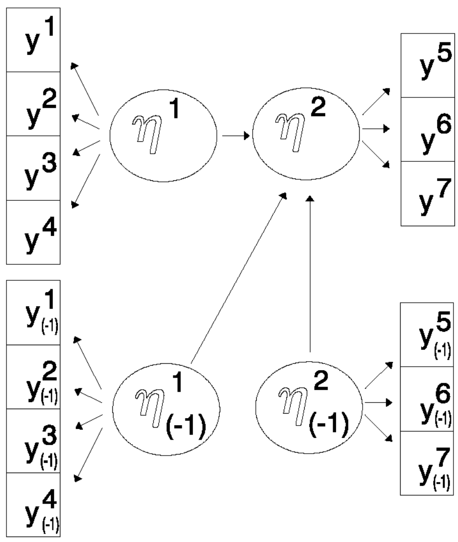

Fig.1: Four block path model for a two LV model

Here $\mathbf{y}^{m}_{(-1)}$, $\boldsymbol{\eta}^{k}_{(-1)}$ denote the lagged data vectors of the variables $\mathbf{y}^{m}$ und $\boldsymbol{\eta}^{k}$ respectively. This path model has the advantage that the classical program LVPLS by Lohmöller (1984) can be used immediately because this model is formally equivalent to a static 4-block path model. Hence the estimation gives different weights for the lagged and unlagged blocks of

MVs, because they are treated as mutually independent variables. This violates the logic of a dynamic model and fails the aim of estimating time constant weight vectors for each LV.

The path model in fig. 2 provides a more consistent solution corresponding to the algorithm presented in this paper. The dotted arrows denote lagged relationships. It obviously avoids the problem of dual weight matrices. The PLS estimation procedure for dynamic path models with latent variables (DPLS) would give definite weight estimations for both latent variables.

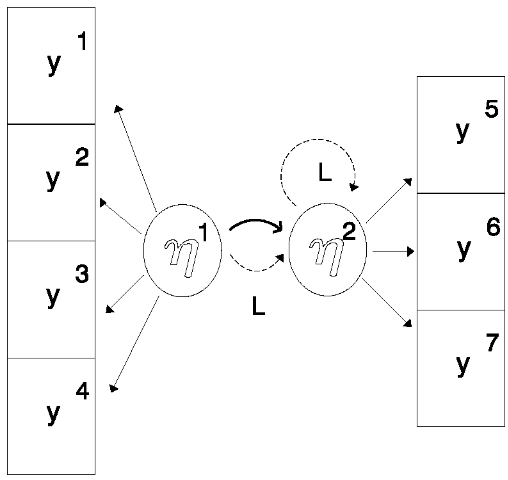

Fig. 2: Two block path model for a two LV dynamic model

The vectors and matrices of the general model (11), (12) have in this particular case the following shape:

the data matrices and variables $\boldsymbol{y}_t = \begin{bmatrix} y_t^1 \\ y_t^2 \\ \vdots \\ y_t^7 \end{bmatrix}$ $\eta_t = \begin{bmatrix} \eta_t^1 \\ \eta_t^2 \end{bmatrix}$ $\boldsymbol{Y} = \lfloor \boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_T \rfloor$ $\boldsymbol{H} = \lfloor \eta_1, \eta_2, \dots, \eta_T \rfloor$, the coefficient matrices $\boldsymbol{W} = \begin{bmatrix} \omega_{11} & 0 \\ \omega_{21} & 0 \\ \omega_{31} & 0 \\ \omega_{41} & 0 \\ 0 & \omega_{52} \\ 0 & \omega_{62} \\ 0 & \omega_{72} \end{bmatrix}$ $\quad \boldsymbol{B} = \begin{bmatrix} 0 & 0 \\ b & 0 \end{bmatrix}$ $\quad \boldsymbol{C} = \begin{bmatrix} 0 & 0 \\ c_1 & c_2 \end{bmatrix}$ $\quad \boldsymbol{P} = \begin{bmatrix} p_{11} & 0 \\ p_{21} & 0 \\ p_{31} & 0 \\ p_{41} & 0 \\ 0 & p_{52} \\ 0 & p_{62} \\ 0 & p_{72} \end{bmatrix}$

and the residual matrices $\boldsymbol{ny}_t = \begin{bmatrix} ny_t^1 \\ ny_t^2 \end{bmatrix}$ $\quad \varepsilon_t = \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \\ \vdots \\ \varepsilon_t^7 \end{bmatrix}$.

4. Conclusions

Partial least squares for dynamic path models with latent variables (DPLS) is an promising tool for exploratory analysis of latent variables path models with autoregressive and distributed lag terms. It will be particularly efficient for models with large numbers of manifest variables and only few latent variables. In these cases it provides models much more parsimonious than vector autoregressive models (cf Hendry,Clemens 1994) directly concerning the manifest variables. Cross validation and redundancy measurement are harmonizing instruments for model evaluation. The aims of modelling with DPLS should be data exploration and forecasting. DPLS is not suitable for confirmative or causal analyses of dependencies. Because of the remaining lack of theoretical knowledge about distributional properties, significance tests are at present not yet appropriate to DPLS models. Monte Carlo simulation studies are in preparation.

References:

Aoki, M. (1990): State space modelling of time series. Springer-Verlag, Berlin u.a.

Apel, H./ Lohmöller, J.-B. (1992): Ökonomie und Umweltqualität; in: Hildebrandt/ Rudinger/Hildebrandt (Hrsg.): Kausalanalysen in der Umweltforschung, Stuttgart, p. 73-100.

Banerjee, A. / Dolado J. / Galbraith, J.W. / Hendry D.F. (1993): Co-integration, error correction and the econometric Analysis of non-stationary data; Oxford University Press.

Bordignon, S. / Gaetan (1989): Indicatori multipli dell 'occupazi' ne e domanda di Lavoro: alcune analisi empiriche con un modello DYMIMIC; in: Economica & Lavoro 23 (2), p. 19-38.

Hellan, K. / Berntsen H.E. / Borgen O.S. / Martens H. (1991): Recursive algorithm for partial least squares regression; in: Chemometrics and Intelligent Laboratory Systems 14, p. 129-137.

Helland, I.S. (1988): On the structure of partial least squares regression; in: Communs Statist. Simuln 17, p. 581-607.

Helland, I.S. (1990): Partial Least Squares Regression and Statistical Models; in: Scandinavian Journal of Statistics 17, p. 97-114.

Hendry, D.F. / Clemens, M.P. (1994): On a theory of intercept corrections in macroeconometric forecasting; in: Holly S. (Ed.): Money, Inflation and Employment: Essays in Honor of James Ball, Edward Elgar Publishing, Hants.

Hillmer, M. (1993): Kausalanalyse makroökonomischer Zusammenhänge mit latenten Variablen. Physica-Verlag, Heidelberg.

Hotelling, H. (1933): Analysis of a complex of statistical variables into principal components; in: Journal of Educational Psychology 24, p. 417-441, 498-520.

Hotelling, H. (1935): The most predictable criterion; in: Journal of Educational Psychology 26, P. 139-142.

Jong, S.d. (1993): SIMPLS: an alternativ approach to partial least squares regression; in: Chemometrics and Intelligent Laboratory Systems 18, p. 251-263.

Jöreskog, K.G./ Sörbom, D. (1987): LISREL VII Program Manual. International Educational Services, Chicago.

Kelley, T.L. (1935): Essential traits of mental life; in: Harvard Studies in Education 26, p. 146.

Lohmöller, J.-B. (1984): LVPLS 1.6 - Program Manual (Latent Variables Path Analysis with Partial Least Squares Estimation). Zentralarchiv für empirische Sozialforschung, Universität Köln.

Lohmöller, J.-B. (1989): Latent Variable Path Modelling with Partial Least Squares. Heidelberg.

Marengo, E. (1991): A fast method for the calculation of partial least squares coefficients; in: Chemometrics and Intelligent Laboratory Systems 12, p. 117-120.

Mathes, H. (1993): Der PLS-Ansatz für die Analyse von Pfadmodellen; Mathematical Systems in Economics. Anton Hain, Frankfurt/Main.

Otter, P.W. (1992): Dynamic Models with Latent Variables from a System Theoretic Perspective: Theory and Applications; in: Statistica 3, p. 347-364.

PC-ISP (1992): Users Guide and Command Descriptions, Datavision AG, Schweiz.

Phatak, A./ Reilly, P.M./ Penlidis,A. (1992): The Geometry of 2-Block Partial least squares regression; in: Communications in Statistics; A: Theory and Methods 21 (6), p. 1517-1553.

Rönz, B./ Strohe, H.G. (eds.) (1994): Lexikon Statistik. Gabler, Wiesbaden.

Schneeweiß, H. (1993): Consistency at Large in Models with Latent Variables; in: Haagen / Bartholomew / Deistler (eds): Statistical Modelling and Latent Variables, Elsevier, p. 299-320.

Spearman, C. (1927): The Abilities of Man. Macmillan, New York.

Stone, M./ Brooks, R.J. (1990): Continuum Regression: Cross-validated Sequentially Constructed Prediction Embracing Ordinary Least Squares, Partial Least Squares and Principal Components Regression; in: Journal of the Royal Statistical Society - Series B 52; p. 237-269.

Strohe, H. G. (1992): Strukturmodelle mit latenten Variablen - Pfadmodellierung in den Wirtschaftswissenschaften; Wiss. Zeitschrift der Humboldt-Univ., Reihe Geistes- und Sozialwiss. 41, p. 109-120.

Strohe, H.G. (1993): Weiche Modellierung umweltökonomischer Zusammenhänge; in: Allgemeines Statistisches Archiv 77, p. 281-310.

Thurstone, L.L. (1935): Vectors of the Mind. University of Chicago Press, Chicago, IL.

Trivellato, U. / Bordignon, S. / Gaetan, C. (1993): A DYMIMIC model of employment: another look at some issues of formulation, identification and estimation; in: Haagen / Bartholomew / Deistler (eds): Statistical Modelling and Latent Variables, Elsevier, p. 321-340.

Wold, H. (1973): Nonlinear Iterative Partial Least Squares (NIPALS) Modelling - Some Current Developement; in P.R. Krishnajah (Ed.), Multivariate Analysis (Vol. 3, p. 383-407), New York; Academic Press.

Wright, S. (1934): The Method of Path Coefficients; Annals of Mathematical Statistics 5, p. 162-215.