SEM modeling with singular moment matrices

Part III: GLS estimation

Hermann Singer

Diskussionsbeitrag Nr. 491

October 2015

Diskussionsbeiträge der Fakultät für Wirtschaftswissenschaft
der FernUniversität in Hagen
Herausgegeben vom Dekan der Fakultät

SEM modeling with singular moment matrices

Part III: GLS estimation

Hermann Singer
FernUniversität in Hagen *

October 23, 2015

Abstract

We discuss Generalized Least Squares (GLS) and Maximum Likelihood (ML) estimation for structural equations models (SEM), when the sample moment matrices are possibly singular. This occurs, e.g., for panel data when there are more panel waves than independent replications, or for time series data, where the number of time points is large, but only one unit is observed. In preceeding papers, it was shown that ML estimation of the SEM is possible by using a correct gaussian likelihood function. In this article, the usual GLS fit function is modified so that it is defined for singular sample moment matrices.

Key Words: Structural Equation Models (SEM); Panel Data; Generalized Least Squares (GLS) Estimation; Maximum Likelihood (ML) Estimation. Pseudo Maximum Likelihood (PML) Estimation.

1 Introduction

Structural equations models (SEM) are usually estimated using cross sectional or panel data with many independent replications N. Then, the sample moment matrices of the observed data are nonsingular and their inverses can be computed. This is necessary when using the ML or GLS fit functions of well known program packages (e.g. Jöreskog and Sörbom; 2001). In former articles (Singer; 2010, 2012) it was shown that

*Lehrstuhl für angewandte Statistik und Methoden der empirischen Sozialforschung, D-58084 Hagen, Germany, hermann.singer@fernuni-hagen.de

1. ML estimation is also possible for singular sample moment matrices occuring in small samples, and

2. that the results coincide with recursive Kalman filter methods well known in control engineering and econometrics (cf., e.g., Watson and Engle; 1983; Caines; 1988).

In this case, the asymptotics of the estimators (consistency, asymptotic normality etc.) are not considered over the cross sectional sample size N but as a function of the number of time points or panel waves T.

More generally, one can consider the parameter estimators as a function of the dimension K of the indicators $y_n, n = 1, ..., N$ for fixed N and a fixed number u of different parameters. For example, one may formulate idiographic models for a single person, which are estimated on several time points, but without the regular structure of an ARIMAX time series model. Such individual causal structures may be used, for example, in psychotherapy research or homeopathy, where standard models are not flexible enough.

In this paper, the problem is discussed in the context of least squares estimation (cf. Browne; 1974). Here, a positive definite weight matrix is used in the fit function, usually the sample covariance matrix. Since this may be singular when using too less cross sectional units (for example time series data where $N = 1$), it is proposed to use as weight the theoretical covariance matrix of the manifest variables, evaluated at the current estimate of the parameter vector or at some reference point in parameter space.

GLS estimation is an alternative to Gaussian ML estimation, when the distribution of the data strongly deviates from normality. Alternatively, one can consider pseudo maximum likelihood (PML) estimation (cf. Gourieroux et al.; 1984; Arminger and Schoenberg; 1989), where a pseudo-likelihood is used which does not coincide with the true density function of the data. In this context, the gaussian pseudo-likelihood function can be considered as a member of the quadratic exponential family.

In section 2, the SEM model is defined and the likelihood function is given in several forms. Then, in section 3, the objective function for generalized least squares is derived. We obtain well known results and a generalization for varying intercepts. Then, the aymptotic standard errors for ML, pseudo-ML and GLS estimation are contrasted (section 4). Finally, the different estimation procedures are compared in simulation study for several sample sizes, using gaussian and nongaussian data with leptokurtic error terms.

2 SEM modeling

In the following the SEM model

$$\begin{aligned} \eta_n &= B\eta_n + \Gamma x_n + \zeta_n \quad (1) \\ y_n &= \Lambda\eta_n + \tau x_n + \epsilon_n \quad (2) \end{aligned}$$

$n = 1, \ldots, N$, will be considered. The structural matrices have dimensions $B : P \times P$, $\Gamma : P \times Q$, $\Lambda : K \times P$, $\tau : K \times Q$ and $\zeta_n \sim N(0, \Sigma_\zeta)$, $\epsilon_n \sim N(0, \Sigma_\epsilon)$ are mutually independent normally distributed error terms $\Sigma_\zeta : P \times P$, $\Sigma_\epsilon : K \times K$. We assume that all structural matrices depend on a parameter vector $\psi : u \times 1$, i.e. $\Sigma_\zeta(\psi)$ etc. For example one can specify $\Sigma_\zeta(\psi) = G_\zeta(\psi)G'_\zeta(\psi)$ to obtain a positive semidefinite matrix. The true parameter vector will be denoted as ψ_0.

In the structural and the measurement model, the variables x_n are *deterministic* control variables. They can be used to model intercepts and for dummy coding. Stochastic exogenous variables ξ_n are already included by extending the latent variables $\eta_n \to \{\eta_n, \xi_n\}$. For example, the LISREL model with intercepts is obtained as

$$\begin{aligned} \begin{bmatrix} \eta_n \\ \xi_n \end{bmatrix} &= \begin{bmatrix} B & \Gamma \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \eta_n \\ \xi_n \end{bmatrix} + \begin{bmatrix} \alpha \\ \kappa \end{bmatrix} 1 + \begin{bmatrix} \zeta_n \\ \zeta_n^* \end{bmatrix} \\ \begin{bmatrix} y_n \\ x_n \end{bmatrix} &= \begin{bmatrix} \Lambda_y & 0 \\ 0 & \Lambda_x \end{bmatrix} \begin{bmatrix} \eta_n \\ \xi_n \end{bmatrix} + \begin{bmatrix} \tau_y \\ \tau_x \end{bmatrix} 1 + \begin{bmatrix} \epsilon_n \\ \delta_n \end{bmatrix} \\ \mathrm{Var}(\begin{bmatrix} \zeta_n \\ \zeta_n^* \end{bmatrix}) &= \begin{bmatrix} \Psi & 0 \\ 0 & \Phi \end{bmatrix} \\ \mathrm{Var}(\begin{bmatrix} \epsilon_n \\ \delta_n \end{bmatrix}) &= \begin{bmatrix} \Sigma_\epsilon & 0 \\ 0 & \Sigma_\delta \end{bmatrix}. \end{aligned}$$

Since the error vectors are normally distributed, the indicators y_n in the measurement model (2) are distributed as $N(\mu_n, \Sigma)$, where

$$\begin{aligned} \eta_n &= B_1(\Gamma x_n + \zeta_n) \\ E[\eta_n] &= B_1\Gamma x_n \\ \mathrm{Var}(\eta_n) &= B_1\Sigma_\zeta B_1' \end{aligned}$$

$$\begin{aligned} E[y_n] &:= \mu_n(\psi) = \Lambda E[\eta_n] + \tau x_n = [\Lambda B_1\Gamma + \tau]x_n := C(\psi)x_n \\ \mathrm{Var}(y_n) &:= \Sigma(\psi) = \Lambda \mathrm{Var}(\eta_n)\Lambda' + \Sigma_\epsilon = \Lambda B_1\Sigma_\zeta B_1'\Lambda' + \Sigma_\epsilon. \end{aligned}$$

In the equations above, it is assumed that $B_1 := (I - B)^{-1}$ exists. In short

form one can write the SEM as a regression equation[1]

$$\begin{aligned} y_n &= \mu_n(\psi) + \nu_n = C(\psi)x_n + \nu_n \quad (3) \\ \nu_n &\sim N(0, \Sigma(\psi)). \quad (4) \end{aligned}$$

Thus, the log likelihood function for the N observations $\{y_n, x_n\}$ is

$$l(\psi) = -\tfrac{N}{2}\left(\log|\Sigma| + \mathrm{tr}\left[\Sigma^{-1}\ \tfrac{1}{N}\sum_n (y_n - \mu_n)(y_n - \mu_n)'\right]\right). \quad (5)$$

Inserting μ_n (eqn. 3) and using the data matrices $Y' = [y_1, ..., y_N] : K \times N$, $X' = [x_1, ..., x_N] : Q \times N$, the log likelihood can be written as

$$l = -\tfrac{N}{2}\left(\log|\Sigma| + \mathrm{tr}\left[\Sigma^{-1}(M_y + CM_xC' - M_{yx}C' - CM_{xy})\right]\right), \quad (6)$$

with the empirical moment matrices $M_y = N^{-1}Y'Y : K\times K$, $M_x = N^{-1}X'X : Q \times Q$, $M_{yx} = N^{-1}Y'X : K \times Q$.

In order to find the relation to GLS estimation, one can insert the sample covariance matrix $S = \frac{1}{N}\sum_n (y_n - \bar{y})(y_n - \bar{y})'$ in (5) which yields the form (for the case $\mu_n = \mu$)

$$l = -\tfrac{N}{2}\left(\log|\Sigma| + \mathrm{tr}\left\{\Sigma^{-1}\ [S + (\bar{y} - \mu)(\bar{y} - \mu)']\right\}\right). \quad (7)$$

More generally one finds

$$\begin{aligned} l &= -\tfrac{N}{2}\Big(\log|\Sigma| + \mathrm{tr}\big\{\Sigma^{-1}[S + \frac{1}{N}\sum_n (\bar{y} - \mu_n)(\bar{y} - \mu_n)' \\ &\qquad + (y_n - \bar{y})(\bar{y} - \mu_n)' + (\bar{y} - \mu_n)(y_n - \bar{y})']\big\}\Big). \quad (8) \end{aligned}$$

For $\mu_n = \mu$, the mixed product terms in the second line are null and one recovers (7).

In contrast to ML estimation, in least squares estimation no probability distribution of the data is assumed. Thus one may define the equation errors as $\zeta_n \sim (0, \Sigma_\zeta)$, $\epsilon_n \sim (0, \Sigma_\epsilon)$ without normality assumption but retains the correct specification of the first and second moments μ_n and Σ. As will be shown in the next section, the GLS fit function for the model without intercepts is given in the usual form as

$$F = \tfrac{N}{2}\ \mathrm{tr}\left[(\Sigma - S)V\right]^2, \quad (9)$$

[1]The dependence of $\mu_n(\psi)$ and $\Sigma(\psi)$ will be displayed only when necessary.

where the weight matrix $V = S^{-1}$ is the inverse sample covariance matrix of y_n. The so defined GLS fitting function requires the positive definiteness (and thus nonsingularity) of S.

In cases of singular (or nearly singular) S, it is proposed to use the variable $V = \Sigma^{-1}(\psi)$ or other nonsingular constant matrices as weight function.

In contrast, the likelihood function (7) is well defined for singular S ($N \leq K$), since no log determinants of the sample moment matrices are involved, as is suggested by the ML fitting function of LISREL (cf. LISREL 8 reference guide, p. 21, eqns. 1.14, 1.15, p. 298, eqn. 10.8; Jöreskog and Sörbom 2001). In Browne (1974), this is called a Wishart likelihood function. The covariance matrix $\Sigma(\psi)$ (eqn. 3) of the indicators y_n must be nonsingular, however.[2]

In the case of small N, in extreme form $N = 1$, the asymptotics of the estimators must be considered as a function of dimension $K = \dim(y_n)$. For example, in time series analysis, we have $y_n = \{y_{n0}, ..., y_{nT}\}, n = N = 1$, so that $K = T + 1$ is the number of time points or panel waves.

If the error terms are not normally distributed, the likelihood (5) can be considered as a pseudo likelihood (cf. Gourieroux et al.; 1984; Arminger and Schoenberg; 1989) with correct first and second moments. It yields consistent estimates, but requires corrections in the asymptotic standard errors (see section 4).

3 Least Squares Estimation

We propose the general least squares criterion

$$\begin{aligned} F(\psi) &= (s-\sigma)'W^{-1}(s-\sigma) + N^{-1}\sum_n (\bar{y}-\mu_n)'W_1^{-1}(\bar{y}-\mu_n) \qquad (10) \\ &= F_2 + F_1, \end{aligned}$$

where $s = \text{rows}(S) = s_{ij}, i \leq j \leq K : \tilde{K} \times 1, \tilde{K} := \frac{1}{2}K(K+1)$ is the vectorized upper triangle of matrix $S : K \times K$ and $\sigma(\psi) = \text{rows}(\Sigma(\psi))$. One can write $\text{rows}(S) = D^+\text{row}(S)$ with the duplication matrix $D : K^2 \times \tilde{K}$ defined as $\text{row}(S) = D\ \text{rows}(S)$ and with $D^+ : \tilde{K} \times K^2$ as its pseudoinverse (Magnus and Neudecker; 1999, ch. 3). Furthermore, $\text{row}(S) = s_{ij}, i, j \leq K$ is the row-wise vectorized matrix $S : K \times K$.

Usually, the weight matrix W is chosen as the covariance matrix of the sample covariances s_{ij}, i.e. $W_{ij,kl} = \text{Cov}(s_{ij}, s_{kl}); i \leq j, k \leq l$. This choice is called a *correct weight matrix* by Jöreskog (1990).

[2]Otherwise the singular normal distribution can be used (Mardia et al.; 1979, p. 41). This case occurs in the presence of restrictions between the components of y_n.

One can write

$$W = \mathrm{Cov}(s,s) = D^{+}\ \mathrm{Cov}(\mathrm{row}(S),\mathrm{row}(S))\ D^{+'}. \tag{11}$$

Inserting the sample covariance matrix

$$s_{ij} = N^{-1}\sum_{n} y_{ni}y_{nj} - \bar{y}_i\bar{y}_j := m_{ij} - m_i m_j \tag{12}$$

one obtains

$$\begin{aligned}\mathrm{Cov}(s_{ij},s_{kl}) &= \mathrm{Cov}(m_{ij},m_{kl}) - \mathrm{Cov}(m_{ij},m_k m_l)\\ &- \mathrm{Cov}(m_{kl},m_i m_j) + \mathrm{Cov}(m_i m_j, m_k m_l).\end{aligned} \tag{13}$$

The first term is given by

$$\mathrm{Cov}(m_{ij},m_{kl}) = N^{-2}\sum_{n}\mathrm{Cov}(y_{ni}y_{nj},y_{nk}y_{nl})$$

and inserting $y_n = \mu_n + \nu_n$ (equation 3) one gets (setting $\mu_{nik} = \mu_{ni}\mu_{nk}$)

$$\begin{aligned}\mathrm{Cov}(y_{ni}y_{nj},y_{nk}y_{nl}) &= \mu_{nik}\sigma_{jl} + \mu_{nil}\sigma_{jk} + \mu_{njk}\sigma_{il} + \mu_{njl}\sigma_{ik}\\ &+ \sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk}.\end{aligned}$$

In deriving this equation, it was assumed that the 4th moments can be written using gaussian error terms as

$$E[\nu_{ni}\nu_{nj}\nu_{nk}\nu_{nl}] = \sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk} + \sigma_{ij}\sigma_{kl}.$$

Otherwise, one has to insert $\mathrm{Cov}(\nu_{ni}\nu_{nj},\nu_{nk}\nu_{nl}) = E[\nu_{ni}\nu_{nj}\nu_{nk}\nu_{nl}] - \sigma_{ij}\sigma_{kl}$ (see Browne; 1984).

Computing the other terms in (13) in an analogous way, the desired weight matrix in (11) is obtained in symbolic form as

$$\mathrm{Cov}(\mathrm{row}(S),\mathrm{row}(S)) = c\cdot[(4)(\overline{\mu\mu'}\otimes\Sigma) + (2)(\Sigma\otimes\Sigma)] \tag{14}$$

where the number in parantheses denotes the possible permutations[3] of the indices i,j,k,l and $\overline{\mu\mu'} = CM_xC', M_x = N^{-1}\sum x_n x_n'$. The factor c is given by $c = N^{-1}(1-2N^{-1}+N^{-2}) \approx N^{-1}$ in large samples.

[3] for example $\sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk} = (2)(\Sigma\otimes\Sigma)$.

Now, the covariance matrix (14) is multiplied in (11) from left and right by the pseudoinverse of the duplication matrix, which leads to the simplification

$$\begin{aligned} W &= c\, D^{+}\, [4(\overline{\mu\mu'} \otimes \Sigma) + 2(\Sigma \otimes \Sigma)]\, D^{+'} \\ &= 2c\, D^{+}\, [(2\overline{\mu\mu'} + \Sigma) \otimes \Sigma]\, D^{+'}. \end{aligned} \tag{15}$$

There is an inversion theorem of the form

$$[D^{+}\, (A \otimes A)\, D^{+'}]^{-1} = D'\, (A^{-1} \otimes A^{-1})\, D \tag{16}$$

(Browne; 1974; Magnus and Neudecker; 1999, see also appendix), so that the weight matrix can be written as (setting Σ to the true value Σ_0)

$$W^{-1} = \tfrac{N}{2}\, D'\, (\Sigma_0^{-1} \otimes \Sigma_0^{-1})\, D \tag{17}$$

in the special case of vanishing intercepts $\overline{\mu\mu'} = 0$. This is much more efficient than the direct inversion of the matrix $W : \tilde{K} \times \tilde{K}$. Then, one can write

$$\begin{aligned} F_2(\Sigma_0) &= (s-\sigma)' W^{-1} (s-\sigma) \\ &= \tfrac{N}{2} (s-\sigma)'\, D'\, (\Sigma_0^{-1} \otimes \Sigma_0^{-1})\, D\, (s-\sigma) \\ &= \tfrac{N}{2}\, \mathrm{row}'(S-\Sigma)\, (\Sigma_0^{-1} \otimes \Sigma_0^{-1})\, \mathrm{row}(S-\Sigma) \\ &= \tfrac{N}{2}\, \mathrm{tr}[(S-\Sigma)\Sigma_0^{-1}]^2. \end{aligned} \tag{18}$$

Here we used the formula $\mathrm{tr}[ABCD] = \mathrm{row}'(A)(D' \otimes B)\mathrm{row}(C')$ (see appendix). Usually, the unknown Σ_0 in the weight matrix is replaced by the estimate S and one obtains

$$F_2(S) = \tfrac{N}{2}\, \mathrm{tr}[(S - \Sigma(\psi))S^{-1}]^2. \tag{19}$$

Thus we have derived the familiar GLS fit function (9). In this paper it is proposed to consider the alternative form with variable weight $\Sigma(\psi)$

$$F_2(\Sigma) = \tfrac{N}{2}\, \mathrm{tr}[(S - \Sigma(\psi))\Sigma^{-1}(\psi)]^2, \tag{20}$$

since $\Sigma(\psi)$ is always nonsingular (cf. Browne; 1974, p. 7, and footnote 2).

In the case with intercepts, i.e. $\overline{\mu\mu'} \neq 0$, one cannot simplify the inverse of weight matrix W (eqn. 15). Alternatively, one could use the form

$$W = \tfrac{2}{N}\, D^{+}\, [(\alpha\, \overline{\mu\mu'} + \Sigma_0) \otimes (\alpha\, \overline{\mu\mu'} + \Sigma_0)]\, D^{+'} \tag{21}$$

with a free parameter α. This leads to the GLS criterion

$$F_2(\alpha, \Sigma_0) = \tfrac{N}{2}\, \mathrm{tr}[(S - \Sigma(\psi))(\alpha\, \overline{\mu\mu'} + \Sigma_0)^{-1}]^2. \tag{22}$$

The weight matrix is nonsingular, since $|\alpha\ \overline{\mu\mu'} + \Sigma_0| \geq |\Sigma_0| > 0$ (Magnus and Neudecker; 1999, p. 21).

Finally, the weight matrix W_1 of the criterion F_1 for the means μ_n (eqn. 10) is given by $\mathrm{Cov}(\bar{y}, \bar{y}) = N^{-1}\Sigma_0$, since $\bar{y} = N^{-1}\sum y_n$. In the case $\mu_n = \mu$, one obtains the familiar form (Browne; 1974; Jöreskog and Sörbom; 2001, p. 298 f.)

$$F_1(\Sigma_0) = (\bar{y} - \mu)' W_1^{-1} (\bar{y} - \mu) = N(\bar{y} - \mu)' \Sigma_0^{-1} (\bar{y} - \mu). \tag{23}$$

Again one can replace the unknown Σ_0 by S or $\Sigma(\psi)$. This form coincides with the last term in the likelihood function (7).

4 Standard errors in GLS and PML estimation

4.1 PML estimation

If the likelihood l is maximized, one can write by Taylor expansion around the true parameter value ψ_0

$$s(\hat{\psi}) \quad = \quad s(\psi_0) + H(\psi_0)(\hat{\psi} - \psi_0) + O(||\hat{\psi} - \psi_0||^2),$$

where the gradient (score) $s(\hat{\psi}) = (\partial l/\partial\psi)(\hat{\psi}) = 0$ at the maximum and the Hessian matrix is $H(\psi_0) = (\partial^2 l/\partial\psi\partial\psi')(\psi_0)$. Thus one obtains

$$\hat{\psi} - \psi_0 \quad \approx \quad -H(\psi_0)^{-1} s(\psi_0). \tag{24}$$

Now, the score is a sum of independent random variables (see 5)

$$s(\psi_0) \quad = \quad \sum_n \partial l_n/\partial\psi_0 = \sum_n s_n(\psi_0), \tag{25}$$

so by the central limit theorem we have the asymptotic distribution

$$s(\psi_0) \quad \sim \quad N(0, F(\psi_0)), \tag{26}$$

where $F(\psi_0) = E_0[s(\psi_0)s'(\psi_0)]$ is the Fisher information matrix and the expectation E_0 is taken with respect to the true distribution. Under the assumption $H(\psi_0) - E_0[H(\psi_0)] \xrightarrow{p} 0$ and using (24, 26), the asymptotic standard errors of $\hat{\psi}$ are given by the 'sandwich' form (Rao 1973, pp. 122, 350, White 1982)

$$\mathrm{Var}(\hat{\psi}) \quad \approx \quad H(\psi_0)^{-1} F(\psi_0) H(\psi_0)^{-1}. \tag{27}$$

The necessity of such a procedure in the SEM context under misspecification was stressed by Arminger and Schoenberg (1989). From (25) one obtains the outer product of gradients (OPG) estimate

$$E[s(\psi_0)s(\psi_0)'] = \sum_n E[s_n(\psi_0)s_n'(\psi_0)] \quad (28)$$

$$\approx \sum_n s_n(\psi_0)s_n'(\psi_0) \quad (29)$$

of the Fisher information matrix. This estimate requires the so called individual likelihood approach (see eqn. 5) and cannot be computed with the moment matrices. If only few cross sectional units are present, the OPG estimate may be singular ($N < \dim(\psi_0)$). To avoid such problems, one can use the Kalman filter to obtain an additional sum over the time points t of the panel waves stacked in the SEM state. This also works for pure time series ($N = 1$). In this case, the conditional scores $s_t = s_{t|t-1} = \partial/\partial\psi\ l(y_t|y_{t-1}, ..., y_0)$ are uncorrelated martingale differences (for details and references, cf. Singer; 2010).

Now, if the model is correctly specified, one has the information matrix identity

$$F(\psi_0) = -E_0[H(\psi_0)]$$

and the asymptotic variance is of the familiar form

$$\mathrm{Var}(\hat{\psi}) \approx F(\psi_0)^{-1} = -E_0[H(\psi_0)]^{-1}.$$

4.2 GLS estimation

In the case of GLS estimation, the criterion F_2 is (see 18)

$$F_2 = (s-\sigma)'W^{-1}(s-\sigma)$$

with gradient

$$g_i(\psi) = 2\sigma_i'\, W^{-1}(\sigma - s),$$

and Hessian

$$H_{ij}(\psi) = 2\sigma_{ij}'\, W^{-1}(\sigma - s) + 2\sigma_i'\, W^{-1}\sigma_j,$$

$\sigma_i := \partial\sigma/\partial\psi_i, \sigma_{ij} := \partial^2\sigma/\partial\psi_i\partial\psi_j$. In large samples, one has the asymptotic result

$$H_{ij}(\psi_0) \xrightarrow{p} 2\sigma_{0i}'\, W^{-1}\sigma_{0j},\ s \xrightarrow{p} \sigma_0 = \sigma(\psi_0),$$

$\sigma_{0i} := (\partial\sigma/\partial\psi_i)(\psi_0)$. Since the covariance matrix of the gradient g_i is

$$\begin{aligned} F_{ij}(\psi) = \mathrm{Cov}(g_i, g_j) &= 4\sigma_i' W^{-1}\mathrm{Cov}(s,s)W^{-1}\sigma_j \\ &= 4\sigma_i' W^{-1}\sigma_j, \end{aligned}$$

one obtains the asymptotic standard errors (see 27)

$$\begin{aligned} \mathrm{Var}(\hat{\psi}) &\approx H(\psi_0)^{-1}F(\psi_0)H(\psi_0)^{-1} \\ &= 2H(\psi_0)^{-1}, \end{aligned}$$

where $H(\psi_0)$ can be written as (Browne; 1974, proposition 2, $\bar{V} = \Sigma_0^{-1}$)

$$H_{ij} = 2\sigma_{0i}' W^{-1}\sigma_{0j} = N\ \mathrm{tr}[\Sigma_{0i}\Sigma_0^{-1}\Sigma_{0j}\Sigma_0^{-1}] \tag{30}$$

(see eqns. 17–18).

5 Simulation study: Continuous time AR(2) panel data

The behaviour of ML and GLS estimators will now be explored for varying sample size N, especially for $N < K$, where the usual weight $V = S^{-1}$ is singular. We use panel data $z_{ni}, n = 1, ..., N; i = 0, ..., T$, which are generated by a vector autoregression with observation error.

5.1 Model specification

The discrete time dynamical state space panel model (vector autoregression VAR(1) with measurement model) is defined by

$$y_{n,i+1} = \alpha_i y_{ni} + \beta_{ni}x_{ni} + u_{ni};\ i = 0, \ldots, T-1 \tag{31}$$

$$z_{ni} = H_i y_{ni} + D_i x_{ni} + \epsilon_{ni};\ i = 0, \ldots, T, \tag{32}$$

$n = 1, ..., N$, with independent Gaussian errors $E[u_{ni}] = 0, \mathrm{Var}(u_{ni}) = \omega_i$, $E[\epsilon_{ni}] = 0, \mathrm{Var}(\epsilon_{ni}) = R_i$. The dimensions of the dynamic structural matrices are $\alpha_i : p \times p$, $\beta_i : p \times q$, $\omega_i : p \times p$, $H_i : k \times p$, $D_i : k \times q$, $R_i : k \times k$. The initial distribution is assumed to be $y_{n0} \sim N(\mu_0, \sigma_0)$ independent of u_{n0} and x_{ni} are deterministic control variables.

This model is very general and permits the treatment of ARIMAX models, dynamic factor analysis, colored noise models etc. (Akaike; 1974; Watson and Engle; 1983; Caines; 1988). All structural matrices depend on a parameter vector ψ.

It can be treated recursively by the Kalman filter or simultaneously by the matrix equation (1–2) where $\eta_n' = [y_{n0}', \ldots, y_{nT}'] : 1 \times (T+1)p$ is the latent SEM state, $\zeta_n' = [\zeta_{n0}', u_{n0}', ..., u_{n,T-1}'] : 1 \times (T+1)p$ is a vector of process errors, $y_n' = [z_{n0}', \ldots, z_{nT}'] : 1 \times (T+1)k$ are the measurements and $x_n' = [1, x_{n0}', \ldots, x_{nT}'] : 1 \times (1 + (T+1)q)$ are (deterministic) exogenous variables.

The structural matrices are given explicitly as

$$B = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 \\ \alpha_0 & 0 & 0 & \ldots & 0 \\ 0 & \alpha_1 & 0 & \ldots & 0 \\ \vdots & 0 & \ddots & 0 & 0 \\ 0 & 0 & \ldots & \alpha_{T-1} & 0 \end{bmatrix}, \; \Sigma_\zeta = \begin{bmatrix} \sigma_0 & 0 & 0 & \ldots & 0 \\ 0 & \omega_0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \vdots & 0 & \ddots & 0 & 0 \\ 0 & 0 & \ldots & 0 & \omega_{T-1} \end{bmatrix}$$

(for the other matrices, cf. Singer; 2010). Solving for η one obtains the solution of the VAR(1)-equation (31) for the time points $t_i, i = 0, ..., T$

$$\eta_n \;\; = \;\; (I - B)^{-1}(\Gamma x_n + \zeta_n). \tag{33}$$

In this equation, the initial condition is represented by $\eta_{n0} = y_{n0} = \mu_0 + \zeta_{n0} \sim N(\mu_0, \sigma_0)$.

We now define a continuous time model which can be written as an exact discrete time model of the form (31). The random oscillator or mathematical pendulum (for details, see Singer; 2012) is defined by the second order stochastic differential equation (SDE)

$$\ddot{y} + \gamma\dot{y} + \omega_0^2 y = bx(t) + g\zeta(t) \tag{34}$$

with the parameters γ = friction, $\omega_0 = 2\pi/T_0$ = angular frequency, T_0 = period of oscillation, g = strength of random force (white noise) $\zeta(t)$ and exogenous controls $x(t)$. The time derivative is denoted as $\dot{y} = dy/dt$.

The pendulum has a continuous-discrete state space representation at the sampling points $t_0, ..., t_T$

$$\begin{aligned} d\begin{bmatrix} y(t) \\ \dot{y}(t) \end{bmatrix} \;\; &:= \;\; \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & -\gamma \end{bmatrix} \begin{bmatrix} y(t) \\ \dot{y}(t) \end{bmatrix} dt + \begin{bmatrix} 0 \\ b \end{bmatrix} x(t)dt + \begin{bmatrix} 0 \\ g \end{bmatrix} dW(t) \\ z_i \;\; &:= \;\; \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} y(t_i) \\ \dot{y}(t_i) \end{bmatrix} + \epsilon_i; \; i = 0, \ldots, T, \end{aligned}$$

$dW = \zeta dt$, where W is the Wiener process. Thus, the SDE of second order can be represented by a first order vector autoregression at the sampling times. Therefore the so called exact discrete model (EDM; Bergstrom 1988) for the sampled states $y_{ni} = [y_n(t_i), \dot{y}_n(t_i)]'$ (a panel of $n = 1, ...N$ oscillators)

has the form

$$
\begin{aligned}
y_{n,i+1} &= A_i^* y_{ni} + B_i^* x_{ni} + u_{ni};\ i = 0, \ldots, T-1 && (35)\\
z_{ni} &= \begin{bmatrix} 1 & 0 \end{bmatrix} y_{ni} + \epsilon_{ni};\ i = 0, \ldots, T, && (36)
\end{aligned}
$$

i.e. (31–32) with the identification $A_i^* = \alpha_i$ etc. In this example only constant controls $x_n(t) = 1$ are considered. The parameter matrices of the EDM are explicitly given as functions of the original model as

$$
\begin{aligned}
A_i^* &= \exp(A\Delta t_i) && (37)\\
B_i^* &= A_i^{-1}(A_i^* - I)B && (38)\\
\mathrm{Var}(u_{ni}) &= \int_0^{\Delta t_i} \exp(As)\Omega \exp(A's)ds, && (39)
\end{aligned}
$$

where

$$
A = \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & -\gamma \end{bmatrix}; B = \begin{bmatrix} 0 \\ b \end{bmatrix}; \Omega = \begin{bmatrix} 0 & 0 \\ 0 & g^2 \end{bmatrix}. \quad (40)
$$

Note that the discrete time error covariance $\mathrm{Var}(u_{ni})$ depends on both the drift and diffusion matrix of (35).

5.2 Simulation study

In the simulation study, the true numerical values were set to $\psi_0 = \{\omega_0^2, \gamma, b, g, \mu_1, \mu_2, \sigma_{11}, \sigma_{12}, \sigma_{22}\} = \{16, 4, 1, 2, 0, 0, 1, 0, 1\}$ where $\mu_k = E[y_{nk}(t_0)], \sigma_{kl} = \mathrm{Cov}(y_{nk}(t_0), y_{nl}(t_0))$ are the parameters of the initial condition. The measurement error variance was set to $R = \mathrm{Var}(\epsilon_{ni}) = 0.01$. The states $y_n(t)$ are assumed to be measured at times $t \in \{0, 0.5, 1, ..., 5.5, 6\}$, i.e. one has $T + 1 = 13$ time points and a regular sampling interval of $\Delta t = 0.5$.

In the simulation study, we consider sample sizes of different order, ranging from $N = 500$, to $N = 1$. The estimation procedure was repeated $M = 100$ times. Since the measured SEM state $y_n = [z_{n0}', ..., z_{nT}']'$ has dimension 13, there may result singular moment matrices $M_y = N^{-1} \sum y_n y_n'$. As shown in former work (Singer; 2010, 2012), the maximum likelihood estimator for the SEM is well defined even for $N = 1$ and coincides with the recursive Kalman filter approach. The usual GLS estimator (9) with $V = S^{-1}$ is not defined for $N < K = 13$ and will be replaced by the choice $V = \Sigma^{-1}(\psi)$ and other nonsingular weight matrices.

Maximization of the likelihood function or minimization of the GLS criterion was achieved using a quasi Newton algorithm with BFGS secant updates (Dennis Jr. and Schnabel; 1983). The iterations in the BFGS algorithm were stopped if both the gradient $||s_k|| < \epsilon$ and the step $||\psi_{k+1} - \psi_k|| < \epsilon$ with $\epsilon = 10^{-4}$.

5.2.1 Gaussian errors

In this section we consider gaussian errors u_{ni} and ϵ_{ni}, so that the likelihood function (6) is correctly specified. In this case, the ML method is expected to give the best results and the GLS method should be equivalent in large samples $N \to \infty$.

Indeed, table 1 indicates the equivalence of ML and GLS in large samples ($N = 500$). Both standard deviations (sd) and bias are small and the methods perform similarly.

Using smaller sample sizes $N = 50$ and $N = 15$, the performance of GLS, especially with weight matrix $V = S^{-1}$ is degraded (tables 2, 3). At $N = 15$, the sample covariance matrix $S : 13 \times 13$ is almost singular, and only in $M' = 15$ of $M = 100$ samples the optimization algorithm converged. In contrast, modified GLS with $V = \Sigma^{-1}$ performs only slightly worse than ML.

Using just one panel unit ($N = 1$), only maximum likelihood (and GLS with weights $\Sigma(\hat{\psi}_{ML})$, see below) leads to satisfying results (table 4). As shown in Singer (2010), one gets the same likelihood as in the Kalman filtering approach. Note that the parameters of the initial condition $y_{n0} = [y_n(t_0), \dot{y}_n(t_0)]'$ cannot be estimated with $N = 1$, thus they were set to the true values. The modified GLS procedure with $V = \Sigma^{-1}(\psi)$ did not converge and was stopped after 200 iterations.

The problems occur, because the proposed weight matrix depends on the parameter vector, so that the modified GLS criterion

$$F_2(\Sigma) \;=\; \tfrac{1}{2}\,\mathrm{tr}[(S-\Sigma)\Sigma^{-1}]^2 = -\tfrac{1}{2}\,\mathrm{tr}[I] = const. \tag{41}$$

is constant for $N = 1, S = 0$ and the modified mean criterion

$$F_1(\Sigma) \;=\; (y-\mu)'\Sigma^{-1}(y-\mu), \tag{42}$$

is only part of $(-2\times)$ the likelihood function

$$l = -\tfrac{1}{2}\left(\log|\Sigma| + \mathrm{tr}\left[\Sigma^{-1}\,(y-\mu)(y-\mu)'\right]\right), \tag{43}$$

but without jacobian term. This leads to incorrect inferences, however, since the gradients of the GLS and the ML criterion do not coincide. One should use, as shown below, a constant weight matrix in F_2, e.g. $V = \Sigma^{-1}(\psi_*)$ where ψ_* is an appropriate value. In the simulation, we used the true parameter vector ψ_0. The results are displayed in table 4.

According to the derivation in Browne (1974, Proposition 6, p. 13), the ML and the GLS criterion are equivalent, if the weight matrix is chosen

as $V = \Sigma(\hat{\psi}_{ML})^{-1}$ and $\hat{\psi}_{ML}$ is the maximum likelihood estimator. This corresponds to using $\psi_* = \hat{\psi}_{ML}$. More precisely, the gradient

$$\begin{aligned} \partial F_2(V)/\partial\psi_i &= \tfrac{1}{2}\,\partial/\partial\psi_i\,\mathrm{tr}[(S-\Sigma)V]^2 \\ &= \mathrm{tr}[V(\Sigma - S)V\Sigma_i] \end{aligned} \tag{44}$$

$\Sigma_i := (\partial\Sigma/\partial\psi_i)$, evaluated at $V = \Sigma(\hat{\psi}_{ML})^{-1}$ coincides with the gradient (score function) of likelihood (7), since

$$\begin{aligned} \partial/\partial\psi_i\,(\log|\Sigma| + \mathrm{tr}[\Sigma^{-1}S]) &= \mathrm{tr}[\Sigma^{-1}\Sigma_i] - \mathrm{tr}[\Sigma^{-1}\Sigma_i\Sigma^{-1}S] \\ &= \mathrm{tr}[\Sigma^{-1}(\Sigma - S)\Sigma^{-1}\Sigma_i]. \end{aligned} \tag{45}$$

This establishes the equivalence of ML and GLS in large samples. On the other hand, the gradient of the modified criterion

$$\begin{aligned} \partial F_2(\Sigma)/\partial\psi_i &= \partial/\partial\psi_i\,\tfrac{1}{2}\mathrm{tr}[(S-\Sigma)\Sigma^{-1}]^2 \\ &= \mathrm{tr}[(S-\Sigma)\Sigma^{-1}S(\Sigma^{-1})_i] \\ &= \mathrm{tr}[\Sigma^{-1}(\Sigma - S)\Sigma^{-1}S\Sigma^{-1}\Sigma_i] \end{aligned} \tag{46}$$

is different, since here $V = \Sigma^{-1}(\psi)$ is an explicit function of ψ. For $N = 1$ ($S = 0$), this gradient is even zero. Only in large samples, the expressions are equivalent, because $S \to \Sigma_0$ and $\Sigma(\hat{\psi}_{GLS}) \to \Sigma_0$.

Since the expressions $F_2(V) = \frac{1}{2}\mathrm{tr}[(S-\Sigma)V]^2$ and $\log|\Sigma| + \mathrm{tr}[\Sigma^{-1}S]$ are equivalent one must consider the criterion

$$F_1(\Sigma) = [\bar{y} - \mu(\psi)]'\Sigma^{-1}(\psi)[\bar{y} - \mu(\psi)] \tag{47}$$

for the mean part, as suggested by equation (7). Using a parameter independent $\Sigma^{-1}(\psi_*)$ in F_1 instead leads to incorrect inferences.

Example A simple example will clarify the issue. Assume that $N = 1$ and consider the AR(1)-time series $y_{t+1} = \phi y_t + \sigma\epsilon_t; t = 0, ..., T-1; y_0 \sim N(0, \sigma_0^2)$. This can be represented by the SEM model $\eta = B\eta + \zeta; \eta = [y_0, ..., y_T]'$, $E[\eta] = 0$ with structural matrices

$$B = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ \phi & 0 & 0 & \dots & 0 \\ 0 & \phi & 0 & \dots & 0 \\ \vdots & 0 & \ddots & 0 & 0 \\ 0 & 0 & \dots & \phi & 0 \end{bmatrix},\ \Sigma_\zeta = \begin{bmatrix} \sigma_0^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma^2 & 0 & \dots & 0 \\ 0 & 0 & \sigma^2 & \dots & 0 \\ \vdots & 0 & \ddots & 0 & 0 \\ 0 & 0 & \dots & 0 & \sigma^2 \end{bmatrix}.$$

Thus, the precision matrix of the indicators is $\Sigma^{-1} = (I-B)'\Sigma_\zeta^{-1}(I-B)$ and the mean $\mu = 0$. Therefore, since $S = 0$, the likelihood function is

$$\begin{aligned} l &= -\tfrac{1}{2}\left(\log|\Sigma| + \operatorname{tr}\left[\Sigma^{-1}yy'\right]\right) \\ &= -\tfrac{1}{2}\left(T\log\sigma^2 + \operatorname{tr}\left[y'(I-B)'\Sigma_\zeta^{-1}(I-B)y\right]\right) \\ &= -\tfrac{1}{2}\left(T\log\sigma^2 + \sum_{t=0}^{T-1}(y_{t+1}-\phi y_t)^2/\sigma^2 + y_0^2/\sigma_0^2\right). \end{aligned}$$

The GLS criterion with weight Σ reads

$$\begin{aligned} F_2(\Sigma) &= \tfrac{1}{2}\operatorname{tr}[(S-\Sigma)\Sigma^{-1}]^2 = -\tfrac{1}{2}\operatorname{tr}[I] = const. \\ F_1(\Sigma) &= (y-\mu)'\Sigma^{-1}(y-\mu) \\ &= \sum_{t=0}^{T-1}(y_{t+1}-\phi y_t)^2/\sigma^2 + y_0^2/\sigma_0^2, \end{aligned}$$

and the gradient w.r.t. ϕ and σ^2 is

$$\begin{aligned} F_1(\Sigma)_\phi &= 0 = -\sum_{t=0}^{T-1}(y_{t+1}-\phi y_t)y_t/\sigma^2 \;\Rightarrow \hat\phi = \frac{\sum_{t=0}^{T-1} y_{t+1}y_t}{\sum_{t=0}^{T-1} y_t^2} \\ F_1(\Sigma)_{\sigma^2} &= 0 = -\sum_{t=0}^{T-1}(y_{t+1}-\phi y_t)^2/\sigma^4. \end{aligned}$$

The second equation cannot be fulfilled (an iterative algorithm yields very large values of σ^2 and does not converge). In contrast, the likelihood function gives the ML equation

$$\begin{aligned} l_{\sigma^2} &= 0 = -\tfrac{1}{2}\left(T/\sigma^2 - \sum_{t=0}^{T-1}(y_{t+1}-\phi y_t)^2/\sigma^4\right) \text{ with the familiar solution} \\ \hat\sigma^2 &= T^{-1}\sum_{t=0}^{T-1}(y_{t+1}-\phi y_t)^2. \end{aligned}$$

Therefore, the criterion $F_2(\Sigma)$ must be modified to $F_2(\Sigma_*) = -\frac{1}{2}\operatorname{tr}[\Sigma\Sigma_*^{-1}]^2$ (see (44) with $V = \Sigma_*^{-1}$ and $S = 0$). In an iterative minimization procedure, one could insert $\Sigma_* = \Sigma(\psi_k)$ in the kth iteration (see table 4).

In summary, the maximum likelihood procedure is preferable, since it is well defined for all sample sizes N. In contrast, the GLS procedure does not work, if $V^{-1} = S$ is singular. Using $V = \Sigma^{-1}(\psi)$ instead, one obtains

a criterion which is not equivalent to ML in small samples, unless one substitutes into F_2 the constant matrix $V = \Sigma^{-1}(\hat{\psi}_{ML})$ or $V = \Sigma^{-1}(\psi_*)$ for an appropriety chosen vector. Alternatively, one can insert $V = \Sigma(\psi_k)^{-1}$. The weight in F_1 must be the variable $\Sigma^{-1}(\psi)$, however, in order to retain the consistency property of the GLS estimate. These observations are supported by a look at tables 4 and 8.

5.2.2 Student-t errors

In the case of misspecification, e.g. using nongaussian errors, the likelihood function (6) does not have the correct form. Therefore, GLS estimation should have a better performance in this case. For the simulation, we used equation errors which are Student-t-distributed in order to model leptokurtosis. Random numbers with $\nu = 5$ degrees of freedom were used, leading to a kurtosis of $k = m_4/m_2^2 = 3(\nu-2)/(\nu-4) = 9$. This occurs in applications, for example, in modeling returns of financial data, e.g. stock prices.

In contrast to the expectation, GLS did not perform better than ML. Again, for large samples, the methods perform about the same, but in smaller samples, GLS shows degraded performance.

As already mentioned, the ML estimator under misspecification is called a pseudo-ML (PML) estimator, which is still consistent, but displays different asympotic standard errors (cf. White; 1982; Gourieroux et al.; 1984; Arminger and Schoenberg; 1989; Wefelmeyer; 1996). As explained in section 4, one must use a 'sandwich' form in this case, which reflects the difference in the expected Hessian $-H = -E[\partial^2 l/\partial\psi\partial\psi']$ and the Fisher information matrix $F = E[\partial l/\partial\psi \; \partial l/\partial\psi']$ under mispecification.

5.3 Discussion

Generally, the parameter estimates using misspecified models display larger root mean square errors in comparison to the gaussian data. The differences between pseudo maximimum likelihood and GLS are not very pronounced, however. In large to small samples ($N = 500, 50, 15$) the GLS estimator with variable weight $V = \Sigma(\psi)^{-1}$ performs better than the conventional estimator with constant weight $V = S^{-1}$. This is due to the fact that $\Sigma(\psi)$ is positive semidefinite, but S may become nearly singular. However, in $N = 1$ estimation, where $S = 0$, the proposed modified GLS estimator with $V = \Sigma(\psi)^{-1}$ also displays problems. In section 5.2.1, it was shown that $F_2(\Sigma(\psi)) = \frac{1}{2}\mathrm{tr}[(S - \Sigma)\Sigma^{-1}]$ of the GLS criterion is constant, and the remaining part $F_1 = (\bar{y} - \mu)'\Sigma^{-1}(\bar{y} - \mu)$ yields biased estimates. Still, the ML estimator works well, since it coincides with the recursive prediction error decomposition computed by the Kalman filter. The GLS criterion must

be taken as $F_2(V) = \frac{1}{2}\text{tr}[(S - \Sigma(\psi))V]^2, V = \Sigma^{-1}(\psi_*)$ and $F_1(\Sigma) = [\bar{y} - \mu(\psi)]'\Sigma^{-1}(\psi)[\bar{y} - \mu(\psi)]$.

6 Conclusion

In large samples, GLS and ML estimation perform similarly, but the modified GLS approach with variable weight matrix $\Sigma(\psi)$ instead of S is a good alternative when S becomes nearly singular. Both GLS approaches do not work for $N = 1$, since here $S = 0$ and the modified GLS approach yields biased estimates. As a remedy, one can insert weight matrices depending on a reference point in parameter space. A further alternative is ULS estimation (weight $V = I$). In any case, ML estimation (and pseudo ML under misspecification) works well, coincides with the recursive Kalman filter estimates, and is thus recommended.

Appendix

There is an inversion theorem for Kronecker products of the form (Browne; 1974; Magnus and Neudecker; 1999)

$$[D^+ (A \otimes A) D^{+'}]^{-1} = D' (A^{-1} \otimes A^{-1}) D, \tag{48}$$

where D is the so called duplication matrix with the property

$$\text{row}(S) = D \, \text{rows}(S), \tag{49}$$

and $\text{rows}(S) = s_{ij}, i \leq j \leq k : \tilde{k} \times 1, \tilde{k} := \frac{1}{2}k(k+1)$ is the vectorized upper triangle of matrix $S : k \times k$ and $\text{row}(S) = s_{ij}, i, j \leq k : k^2 \times 1$ is the row wise vectorized matrix S.[4]

One can solve for

$$\text{rows}(S) = D^+\text{row}(S) \tag{50}$$

where $D^+ = (D'D)^{-1}D' : \tilde{k} \times k^2$ is the pseudoinverse of D (Magnus and Neudecker; 1999, ch. 3). Equation (48) can be proved by the property

$$DD^+ (A \otimes A) = (A \otimes A)DD^+ \tag{51}$$

[4]In the main text, we used $k = K$, but here the dimension could be confounded with the commutation matrix K.

of the projection matrix $N = DD^+ = D(D'D)^{-1}D', N^2 = N$, since

$$N = \tfrac{1}{2}(I + K) \tag{52}$$

(see below) and K is the so called commutation matrix with the properties

$$\begin{aligned} K \text{ row}(A') &= \text{row}(A) && (53)\\ K(A \otimes B) &= (B \otimes A)K. && (54) \end{aligned}$$

The last formula follows from

$$\text{row}(ABC) = (A \otimes C')\text{row}(B) \tag{55}$$

and K row$(ABC) = K(A \otimes C')$row$(B) =$ row$(C'B'A') = (C' \otimes A)K$ row(B). Furthermore, one has $K = K' = K^{-1}$. This implies $K^2 = KK^{-1} = I$.

This shows, that $N = \frac{1}{2}(I + K)$ fulfils

$$N\ (A \otimes B) = \tfrac{1}{2}[(A \otimes B) + (B \otimes A)K] \tag{56}$$

and thus $N(A \otimes A) = (A \otimes A)N$.

It remains to show that $N = \frac{1}{2}(I+K)$. First, we have the projection property $N^2 = \frac{1}{4}(I+2K+K^2) = N$, since $K^2 = I$. Furthermore $ND = \frac{1}{2}(D+KD) = D$, since $KD = D$. We also have rank$(N) =$ tr$[N] = \frac{1}{2}($tr$[I] +$ tr$[K]) = \frac{1}{2}(k^2 + k) = \frac{1}{2}k(k + 1) =$ rank$[D]$, since N is projection matrix and $D : k^2 \times \frac{1}{2}k(k + 1)$. Then, theorem 2.8 of Magnus and Neudecker (1999, p. 35) implies the factorization

$$N = \tfrac{1}{2}(I + K) = DD^+. \tag{57}$$

Finally, we have

$$\text{tr}[AB] = \sum_{ij} a_{ij}b_{ji} = \text{row}'(A)\ \text{row}(B'), \tag{58}$$

$$\text{tr}[ABCD] = \text{row}'(A)\ \text{row}(D'C'B') = \text{row}'(A)(D' \otimes B)\text{row}(C'). \tag{59}$$

and

$$x'Ax = \text{tr}[x'Ax] = \text{tr}[Axx']. \tag{60}$$

References

Akaike, H. (1974). Markovian representation of of stochastic processes and its application to the analysis of autoregressive moving average processes, *Ann. Inst. Stat. Math.* **26**: 363–387.

Arminger, G. and Schoenberg, R. J. (1989). Pseudo maximum likelihood estimation and a test for misspecification in mean and covariance structure models, *Psychometrika* **54**(3): 409–425.

Bergstrom, A. (1988). The history of continuous-time econometric models, *Econometric Theory* **4**: 365–383.

Browne, M. W. (1974). Generalized least squares estimators in the analysis of covariances structures, *South African Statistical Journal* **8**: 1–24.

Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures, *British Journal of Mathematical and Statistical Psychology* **37**(1): 62–83.

Caines, P. (1988). *Linear Stochastic Systems*, Wiley, New York.

Dennis Jr., J. and Schnabel, R. (1983). *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice Hall, Englewood Cliffs.

Gourieroux, C., Monfort, A. and Trognon, A. (1984). Pseudo maximum likelihood methods: Theory, *Econometrica* **52, 3**: 681–700.

Jöreskog, K. and Sörbom, D. (2001). *LISREL 8. User Reference Guide*, Scientific Software International, Lincolnwood, IL.

Jöreskog, K. G. (1990). New developments in LISREL: analysis of ordinal variables using polychoric correlations and weighted least squares, *Quality and Quantity* **24**(4): 387–404.

Magnus, J. R. and Neudecker, H. (1999). *Matrix Differential Calculus*, 2. edn, Wiley, New York.

Mardia, K., Kent, J. and Bibby, J. (1979). *Multivariate Analysis*, Academic Press, London.

Rao, C. (1973). *Linear Statistical Inference and Its Applications*, 2 edn, Wiley, New York.

Singer, H. (2010). SEM Modeling with Singular Moment Matrices. Part I: ML-Estimation of Time Series., *Journal of Mathematical Sociology* **34, 4**: 301–320.

Singer, H. (2012). SEM modeling with singular moment matrices. Part II: ML-Estimation of Sampled Stochastic Differential Equations., *Journal of Mathematical Sociology* **36, 1**: 22–43.

Watson, M. and Engle, R. (1983). Alternative algorithms for the estimation of dynamic factor, mimic and varying coefficient regression models, *Journal of Econometrics* **23**: 385–400.

Wefelmeyer, W. (1996). Quasi-likelihood models and optimal inference, *Annals of Statistics* **24**: 405–422.

White, H. (1982). Maximum likelihood estimation of misspecified models, *Econometrica* **50**: 1–25.

true	mean	sd	bias	RMSE
		ML		
16.	16.0182	0.5163	0.0182	0.5166
4.	4.0065	0.1232	0.0065	0.1234
1.	1.0002	0.0512	0.0002	0.0512
2.	2.0007	0.0666	0.0007	0.0666
0.	−0.0001	0.0475	−0.0001	0.0475
0.	0.0055	0.0856	0.0055	0.0858
1.	0.9952	0.0335	−0.0048	0.0338
0.	−0.0015	0.1419	−0.0015	0.1419
1.	0.9968	0.1378	−0.0032	0.1379
		$GLS, V = \Sigma^{-1}(\psi)$		
16.	16.0725	0.5241	0.0725	0.5291
4.	4.0162	0.1271	0.0162	0.1282
1.	1.0033	0.052	0.0033	0.0521
2.	2.0416	0.0693	0.0416	0.0808
0.	−0.0002	0.0479	−0.0002	0.0479
0.	0.0015	0.0847	0.0015	0.0847
1.	1.0046	0.0345	0.0046	0.0348
0.	0.0106	0.148	0.0106	0.1484
1.	1.004	0.1449	0.004	0.145
		$GLS, V = S^{-1}$		
16.	15.9141	0.5515	−0.0859	0.5582
4.	3.9911	0.1256	−0.0089	0.1259
1.	0.9948	0.0511	−0.0052	0.0514
2.	1.9166	0.0676	−0.0834	0.1074
0.	−0.0004	0.049	−0.0004	0.049
0.	0.007	0.0871	0.007	0.0874
1.	0.9769	0.0338	−0.0231	0.041
0.	−0.019	0.1421	−0.019	0.1433
1.	0.9818	0.1363	−0.0182	0.1375

Table 1: ML and GLS estimates for sample size $N = 500$ in $M = 100$ replications.

true	mean	sd	bias	RMSE
ML				
16.	16.0464	1.6226	0.0464	1.6233
4.	4.0069	0.4138	0.0069	0.4138
1.	0.9876	0.1808	−0.0124	0.1812
2.	1.9905	0.1986	−0.0095	0.1988
0.	0.0162	0.1413	0.0162	0.1422
0.	0.0011	0.301	0.0011	0.301
1.	0.9821	0.1013	−0.0179	0.1029
0.	0.0184	0.532	0.0184	0.5324
1.	0.8687	0.466	−0.1313	0.4842
$GLS, V = \Sigma^{-1}(\psi)$				
16.	16.4227	1.6919	0.4227	1.7439
4.	4.0799	0.4427	0.0799	0.4499
1.	1.0107	0.1892	0.0107	0.1895
2.	2.3693	0.2465	0.3693	0.444
0.	0.0151	0.1425	0.0151	0.1433
0.	0.	0.3139	0.	0.3139
1.	1.0728	0.1221	0.0728	0.1422
0.	0.0783	0.6485	0.0783	0.6532
1.	0.9087	0.5733	−0.0913	0.5805
$GLS, V = S^{-1}$				
16.	14.4569	2.7156	−1.5431	3.1234
4.	3.7141	0.6861	−0.2859	0.7433
1.	0.8967	0.2479	−0.1033	0.2686
2.	1.1099	0.2375	−0.8901	0.9213
0.	0.0112	0.1659	0.0112	0.1663
0.	0.0594	0.3157	0.0594	0.3212
1.	0.692	0.3823	−0.308	0.4909
0.	−0.151	0.6563	−0.151	0.6734
1.	0.5239	0.6302	−0.4761	0.7898

Table 2: ML and GLS estimates for $N = 50$.

true	mean	sd	bias	RMSE
		ML		
16.	15.9613	3.4995	-0.0387	3.4997
4.	4.083	0.7584	0.083	0.763
1.	0.969	0.323	-0.031	0.3245
2.	1.9653	0.3823	-0.0347	0.3838
0.	-0.0367	0.2499	-0.0367	0.2525
0.	-0.0216	0.6107	-0.0216	0.6111
1.	0.9411	0.1771	-0.0589	0.1867
0.	0.0765	1.2136	0.0765	1.2161
1.	0.6612	0.7255	-0.3388	0.8007
		$GLS, V = \Sigma^{-1}(\psi)$		
16.	16.6881	4.3273	0.6881	4.3816
4.	4.0854	0.9078	0.0854	0.9118
1.	1.0167	0.3469	0.0167	0.3473
2.	3.049	0.6645	1.049	1.2418
0.	-0.0355	0.2546	-0.0355	0.2571
0.	-0.0232	0.6899	-0.0232	0.6903
1.	1.2896	0.3896	0.2896	0.4854
0.	0.5239	2.2053	0.5239	2.2667
1.	0.6759	1.037	-0.3241	1.0864
		$GLS, V = S^{-1}$		
16.	14.0127	7.6897	-1.9873	7.9423
4.	1.4276	1.3894	-2.5724	2.9237
1.	0.9725	0.4985	-0.0275	0.4992
2.	0.	0.	$-2.$	2.
0.	0.0781	0.2942	0.0781	0.3044
0.	-0.3707	0.681	-0.3707	0.7753
1.	0.0161	0.1433	-0.9839	0.9943
0.	0.0507	0.201	0.0507	0.2073
1.	0.	0.	$-1.$	1.

Table 3: ML and GLS estimates for $N = 15$. Only $M = 15$ converged samples for $GLS, V = S^{-1}$.

true	mean	sd	bias	RMSE
ML				
16.	18.6439	11.0433	2.6439	11.3553
4.	3.8959	5.6667	−0.1041	5.6677
1.	1.0225	1.4178	0.0225	1.4179
2.	1.4648	1.3966	−0.5352	1.4957
$GLS, V = \Sigma^{-1}(\psi)$				
16.	19.2996	21.925	3.2996	22.1719
4.	2.9023	2.7714	−1.0977	2.9809
1.	1.508	2.7375	0.508	2.7842
2.	192.746	6.8787	190.746	190.87
$GLS, V = \Sigma^{-1}(\psi_0)$				
16.	16.6644	2.762	0.6644	2.8408
4.	4.0813	0.8779	0.0813	0.8816
1.	0.8726	1.071	−0.1274	1.0786
2.	1.9351	0.3679	−0.0649	0.3736
$GLS, V = \Sigma^{-1}(\hat{\psi}_{ML})$				
16.	18.9343	10.8163	2.9343	11.2072
4.	3.7287	2.9453	−0.2713	2.9578
1.	0.8919	1.498	−0.1081	1.5019
2.	1.4114	1.0331	−0.5886	1.189
$GLS, V_k = \Sigma^{-1}(\psi_k)$				
16.	18.568	10.5717	2.568	10.8791
4.	4.3131	5.3442	0.3131	5.3534
1.	1.0006	1.3649	0.0006	1.3649
2.	1.7079	1.204	−0.2921	1.2389
$ULS, V = I$				
16.	20.299	14.3775	4.299	15.0065
4.	3.9438	2.9254	−0.0562	2.9259
1.	0.8881	1.3799	−0.1119	1.3845
2.	0.8135	2.3359	−1.1865	2.62

Table 4: ML and GLS estimates for $N = 1$. No converged samples for $GLS, V = \Sigma^{-1}(\psi)$. For ψ_0, the true value $\{16, 4, 1, 2\}$ was used (see text). For $V_k = \Sigma^{-1}(\psi_k)$, the GLS estimate ψ_k in the kth iteration was used.

true	mean	sd	bias	RMSE
		ML		
16.	16.8766	0.465	0.8766	0.9923
4.	4.1435	0.1168	0.1435	0.1851
1.	1.0536	0.0636	0.0536	0.0832
2.	2.8677	0.0837	0.8677	0.8718
0.	0.0038	0.0449	0.0038	0.0451
0.	−0.0158	0.1236	−0.0158	0.1246
1.	1.2966	0.065	0.2966	0.3037
0.	0.266	0.1822	0.266	0.3224
1.	1.341	0.1868	0.341	0.3888
		$GLS, V = \Sigma^{-1}$		
16.	16.9229	0.4922	0.9229	1.046
4.	4.1458	0.1221	0.1458	0.1902
1.	1.0527	0.0636	0.0527	0.0826
2.	2.9269	0.0951	0.9269	0.9318
0.	0.0074	0.048	0.0074	0.0485
0.	−0.0192	0.1196	−0.0192	0.1211
1.	1.3029	0.0754	0.3029	0.3121
0.	0.2836	0.1919	0.2836	0.3424
1.	1.3149	0.2097	0.3149	0.3783
		$GLS, V = S^{-1}$		
16.	16.8272	0.53	0.8272	0.9824
4.	4.128	0.1286	0.128	0.1815
1.	1.0453	0.0645	0.0453	0.0788
2.	2.7465	0.0916	0.7465	0.7521
0.	0.0072	0.0492	0.0072	0.0497
0.	−0.0131	0.1183	−0.0131	0.1191
1.	1.2627	0.0742	0.2627	0.273
0.	0.2399	0.1852	0.2399	0.3031
1.	1.2074	0.573	0.2074	0.6093

Table 5: Student-t distributed errors: Pseudo-ML and GLS estimates for sample size $N = 500$ in $M = 100$ replications.

true	mean	sd	bias	RMSE
		ML		
16.	16.9316	1.7038	0.9316	1.9419
4.	4.1199	0.3868	0.1199	0.405
1.	1.0665	0.2189	0.0665	0.2288
2.	2.8546	0.2877	0.8546	0.9017
0.	0.0212	0.1996	0.0212	0.2007
0.	−0.0516	0.4135	−0.0516	0.4167
1.	1.2916	0.2997	0.2916	0.4181
0.	0.2916	0.7182	0.2916	0.7752
1.	1.1305	0.837	0.1305	0.8471
		$GLS, V = \Sigma^{-1}$		
16.	17.1389	2.07	1.1389	2.3626
4.	4.1667	0.4975	0.1667	0.5247
1.	1.0778	0.2323	0.0778	0.245
2.	3.3844	0.4237	1.3844	1.4478
0.	0.0152	0.1915	0.0152	0.1921
0.	−0.0589	0.4445	−0.0589	0.4484
1.	1.4106	0.2688	0.4106	0.4908
0.	0.3826	0.8855	0.3826	0.9646
1.	0.9258	1.0952	−0.0742	1.0977
		$GLS, V = S^{-1}$		
16.	16.1292	3.6127	0.1292	3.615
4.	4.0094	0.8591	0.0094	0.8592
1.	1.0167	0.3062	0.0167	0.3067
2.	1.7237	0.3376	−0.2763	0.4363
0.	0.0085	0.2339	0.0085	0.2341
0.	−0.0145	0.4834	−0.0145	0.4836
1.	0.9341	0.4867	−0.0659	0.4912
0.	−0.073	1.2438	−0.073	1.246
1.	1.0072	0.9391	0.0072	0.9391

Table 6: Student-t distributed errors: Pseudo-ML and GLS estimates for sample size $N = 50$ in $M = 100$ replications.

true	mean	sd	bias	RMSE
		ML		
16.	17.3848	3.4681	1.3848	3.7343
4.	4.2247	0.9718	0.2247	0.9974
1.	1.1081	0.4541	0.1081	0.4668
2.	2.8862	0.6801	0.8862	1.1171
0.	0.0305	0.3511	0.0305	0.3524
0.	0.0818	0.6918	0.0818	0.6966
1.	1.2008	0.3076	0.2008	0.3673
0.	0.5644	1.4962	0.5644	1.5991
1.	0.9521	1.0228	−0.0479	1.024
		$GLS, V = \Sigma^{-1}$		
16.	17.8576	4.3553	1.8576	4.7349
4.	4.197	1.0184	0.197	1.0373
1.	1.1264	0.4726	0.1264	0.4892
2.	4.3416	1.0146	2.3416	2.552
0.	0.0388	0.3539	0.0388	0.356
0.	0.1587	0.8673	0.1587	0.8817
1.	1.6433	0.5644	0.6433	0.8558
0.	1.4809	3.173	1.4809	3.5015
1.	1.0858	1.7283	0.0858	1.7305
		$GLS, V = S^{-1}$		
16.	14.8731	6.9453	−1.1269	7.0361
4.	0.649	1.0344	−3.351	3.5071
1.	0.9949	0.8386	−0.0051	0.8386
2.	0.	0.	−2.	2.
0.	−0.0599	0.2675	−0.0599	0.2741
0.	−0.136	0.7198	−0.136	0.7325
1.	0.0408	0.1057	−0.9592	0.965
0.	−0.021	0.2093	−0.021	0.2103
1.	0.	0.	−1.	1.

Table 7: Student-t distributed errors: Pseudo-ML and GLS estimates for sample size $N = 15$ in $M = 100$ replications.

true	mean	sd	bias	RMSE
		ML		
16.	16.1532	9.5657	0.1532	9.5669
4.	2.7765	1.8707	−1.2235	2.2353
1.	1.1478	1.438	0.1478	1.4456
2.	1.8178	1.2978	−0.1822	1.3106
		$GLS, V = \Sigma^{-1}(\psi)$		
16.	6.2775	13.4566	−9.7225	16.6014
4.	0.7697	2.131	−3.2303	3.8699
1.	0.3389	2.2606	−0.6611	2.3553
2.	127.768	73.6055	125.768	145.724
		$GLS, V = \Sigma^{-1}(\psi_0)$		
16.	16.5718	3.4107	0.5718	3.4584
4.	4.0293	1.0183	0.0293	1.0187
1.	1.1026	1.2822	0.1026	1.2863
2.	2.1322	0.5452	0.1322	0.561
		$GLS, V = \Sigma^{-1}(\hat{\psi}_{ML})$		
16.	17.3081	11.0441	1.3081	11.1213
4.	3.1638	2.3342	−0.8362	2.4794
1.	1.1654	1.391	0.1654	1.4008
2.	1.9023	1.3353	−0.0977	1.3388
		$ULS, V = I$		
16.	19.3151	14.0107	3.3151	14.3976
4.	3.6694	2.6024	−0.3306	2.6233
1.	1.0386	1.4952	0.0386	1.4957
2.	1.4278	2.6082	−0.5722	2.6702

Table 8: Student-t distributed errors: Pseudo-ML and GLS estimates for $N = 1$. No converged samples for $GLS, V = \Sigma^{-1}(\psi)$. For ψ_0, the true value $\{16, 4, 1, 2\}$ was used (see text).

Die Diskussionspapiere ab Nr. 183 (1992) bis heute, können Sie im Internet unter http://www.fernuni-hagen.de/wirtschaftswissenschaft/forschung/beitraege.shtml einsehen und zum Teil downloaden.
Ältere Diskussionspapiere selber erhalten Sie nur in den Bibliotheken.

Nr	Jahr	Titel	Autor/en
420	2008	Stockkeeping and controlling under game theoretic aspects	Fandel, Günter Trockel, Jan
421	2008	On Overdissipation of Rents in Contests with Endogenous Intrinsic Motivation	Schlepütz, Volker
422	2008	Maximum Entropy Inference for Mixed Continuous-Discrete Variables	Singer, Hermann
423	2008	Eine Heuristik für das mehrdimensionale Bin Packing Problem	Mack, Daniel Bortfeldt, Andreas
424	2008	Expected A Posteriori Estimation in Financial Applications	Mazzoni, Thomas
425	2008	A Genetic Algorithm for the Two-Dimensional Knapsack Problem with Rectangular Pieces	Bortfeldt, Andreas Winter, Tobias
426	2008	A Tree Search Algorithm for Solving the Container Loading Problem	Fanslau, Tobias Bortfeldt, Andreas
427	2008	Dynamic Effects of Offshoring	Stijepic, Denis Wagner, Helmut
428	2008	Der Einfluss von Kostenabweichungen auf das Nash-Gleichgewicht in einem nicht-kooperativen Disponenten-Controller-Spiel	Fandel, Günter Trockel, Jan
429	2008	Fast Analytic Option Valuation with GARCH	Mazzoni, Thomas
430	2008	Conditional Gauss-Hermite Filtering with Application to Volatility Estimation	Singer, Hermann
431	2008	Web 2.0 auf dem Prüfstand: Zur Bewertung von Internet-Unternehmen	Christian Maaß Gotthard Pietsch
432	2008	Zentralbank-Kommunikation und Finanzstabilität – Eine Bestandsaufnahme	Knütter, Rolf Mohr, Benjamin
433	2008	Globalization and Asset Prices: Which Trade-Offs Do Central Banks Face in Small Open Economies?	Knütter, Rolf Wagner, Helmut
434	2008	International Policy Coordination and Simple Monetary Policy Rules	Berger, Wolfram Wagner, Helmut
435	2009	Matchingprozesse auf beruflichen Teilarbeitsmärkten	Stops, Michael Mazzoni, Thomas
436	2009	Wayfindingprozesse in Parksituationen - eine empirische Analyse	Fließ, Sabine Tetzner, Stefan
437	2009	ENTROPY-DRIVEN PORTFOLIO SELECTION a downside and upside risk framework	Rödder, Wilhelm Gartner, Ivan Ricardo Rudolph, Sandra
438	2009	Consulting Incentives in Contests	Schlepütz, Volker

439	2009	A Genetic Algorithm for a Bi-Objective Winner-Determination Problem in a Transportation-Procurement Auction"	Buer, Tobias Pankratz, Giselher
440	2009	Parallel greedy algorithms for packing unequal spheres into a cuboidal strip or a cuboid	Kubach, Timo Bortfeldt, Andreas Tilli, Thomas Gehring, Hermann
441	2009	SEM modeling with singular moment matrices Part I: ML-Estimation of time series	Singer, Hermann
442	2009	SEM modeling with singular moment matrices Part II: ML-Estimation of sampled stochastic differential equations	Singer, Hermann
443	2009	Konsensuale Effizienzbewertung und -verbesserung – Untersuchungen mittels der Data Envelopment Analysis (DEA)	Rödder, Wilhelm Reucher, Elmar
444	2009	Legal Uncertainty – Is Hamonization of Law the Right Answer? A Short Overview	Wagner, Helmut
445	2009	Fast Continuous-Discrete DAF-Filters	Mazzoni, Thomas
446	2010	Quantitative Evaluierung von Multi-Level Marketingsystemen	Lorenz, Marina Mazzoni, Thomas
447	2010	Quasi-Continuous Maximum Entropy Distribution Approximation with Kernel Density	Mazzoni, Thomas Reucher, Elmar
448	2010	Solving a Bi-Objective Winner Determination Problem in a Transportation Procurement Auction	Buer, Tobias Pankratz, Giselher
449	2010	Are Short Term Stock Asset Returns Predictable? An Extended Empirical Analysis	Mazzoni, Thomas
450	2010	Europäische Gesundheitssysteme im Vergleich – Effizienzmessungen von Akutkrankenhäusern mit DEA –	Reucher, Elmar Sartorius, Frank
451	2010	Patterns in Object-Oriented Analysis	Blaimer, Nicolas Bortfeldt, Andreas Pankratz, Giselher
452	2010	The Kuznets-Kaldor-Puzzle and Neutral Cross-Capital-Intensity Structural Change	Stijepic, Denis Wagner, Helmut
453	2010	Monetary Policy and Boom-Bust Cycles: The Role of Communication	Knütter, Rolf Wagner, Helmut
454	2010	Konsensuale Effizienzbewertung und –verbesserung mittels DEA – Output- vs. Inputorientierung –	Reucher, Elmar Rödder, Wilhelm
455	2010	Consistent Modeling of Risk Averse Behavior with Spectral Risk Measures	Wächter, Hans Peter Mazzoni, Thomas

456	2010	Der virtuelle Peer – Eine Anwendung der DEA zur konsensualen Effizienz-bewertung –	Reucher, Elmar
457	2010	A two-stage packing procedure for a Portuguese trading company	Moura, Ana Bortfeldt, Andreas
458	2010	A tree search algorithm for solving the multi-dimensional strip packing problem with guillotine cutting constraint	Bortfeldt, Andreas Jungmann, Sabine
459	2010	Equity and Efficiency in Regional Public Good Supply with Imperfect Labour Mobility – Horizontal versus Vertical Equalization	Arnold, Volker
460	2010	A hybrid algorithm for the capacitated vehicle routing problem with three-dimensional loading constraints	Bortfeldt, Andreas
461	2010	A tree search procedure for the container relocation problem	Forster, Florian Bortfeldt, Andreas
462	2011	Advanced X-Efficiencies for CCR- and BCC-Modell – Towards Peer-based DEA Controlling	Rödder, Wilhelm Reucher, Elmar
463	2011	The Effects of Central Bank Communication on Financial Stability: A Systematization of the Empirical Evidence	Knütter, Rolf Mohr, Benjamin Wagner, Helmut
464	2011	Lösungskonzepte zur Allokation von Kooperationsvorteilen in der kooperativen Transportdisposition	Strangmeier, Reinhard Fiedler, Matthias
465	2011	Grenzen einer Legitimation staatlicher Maßnahmen gegenüber Kreditinstituten zur Verhinderung von Banken- und Wirtschaftskrisen	Merbecks, Ute
466	2011	Controlling im Stadtmarketing – Eine Analyse des Hagener Schaufensterwettbewerbs 2010	Fließ, Sabine Bauer, Katharina
467	2011	A Structural Approach to Financial Stability: On the Beneficial Role of Regulatory Governance	Mohr, Benjamin Wagner, Helmut
468	2011	Data Envelopment Analysis - Skalenerträge und Kreuzskalenerträge	Wilhelm Rödder Andreas Dellnitz
469	2011	Controlling organisatorischer Entscheidungen: Konzeptionelle Überlegungen	Lindner, Florian Scherm, Ewald
470	2011	Orientierung in Dienstleistungsumgebungen – eine explorative Studie am Beispiel des Flughafen Frankfurt am Main	Fließ, Sabine Colaci, Antje Nesper, Jens

471	2011	Inequality aversion, income skewness and the theory of the welfare state	Weinreich, Daniel
472	2011	A tree search procedure for the container retrieval problem	Forster, Florian Bortfeldt, Andreas
473	2011	A Functional Approach to Pricing Complex Barrier Options	Mazzoni, Thomas
474	2011	Bologna-Prozess und neues Steuerungsmodell – auf Konfrontationskurs mit universitären Identitäten	Jost, Tobias Scherm, Ewald
475	2011	A reduction approach for solving the rectangle packing area minimization problem	Bortfeldt, Andreas
476	2011	Trade and Unemployment with Heterogeneous Firms: How Good Jobs Are Lost	Altenburg, Lutz
477	2012	Structural Change Patterns and Development: China in Comparison	Wagner, Helmut
478	2012	Demografische Risiken – Herausforderungen für das finanzwirtschaftliche Risikomanagement im Rahmen der betrieblichen Altersversorgung	Merbecks, Ute
479	2012	"It's all in the Mix!" – Internalizing Externalities with R&D Subsidies and Environmental Liability	Endres, Alfred Friehe, Tim Rundshagen, Bianca
480	2012	Ökonomische Interpretationen der Skalenvariablen u in der DEA	Dellnitz, Andreas Kleine, Andreas Rödder, Wilhelm
481	2012	Entropiebasierte Analyse von Interaktionen in Sozialen Netzwerken	Rödder, Wilhelm Brenner, Dominic Kulmann, Friedhelm
482	2013	Central Bank Independence and Financial Stability: A Tale of Perfect Harmony?	Berger, Wolfram Kißmer, Friedrich
483	2013	Energy generation with Directed Technical Change	Kollenbach, Gilbert
484	2013	Monetary Policy and Asset Prices: When Cleaning Up Hits the Zero Lower Bound	Berger, Wolfram Kißmer, Friedrich
485	2013	Superknoten in Sozialen Netzwerken – eine entropieoptimale Analyse	Brenner, Dominic, Rödder, Wilhelm, Kulmann, Friedhelm
486	2013	Stimmigkeit von Situation, Organisation und Person: Gestaltungsüberlegungen auf Basis des Informationsverarbeitungsansatzes	Julmi, Christian Lindner, Florian Scherm, Ewald
487	2014	Incentives for Advanced Abatement Technology Under National and International Permit Trading	Endres, Alfred Rundshagen, Bianca

488	2014	Dynamische Effizienzbewertung öffentlicher Dreispartentheater mit der Data Envelopment Analysis	Kleine, Andreas Hoffmann, Steffen
489	2015	Konsensuale Peer-Wahl in der DEA -- Effizienz vs. Skalenertrag	Dellnitz, Andreas Reucher, Elmar
490	2015	Makroprudenzielle Regulierung – eine kurze Einführung und ein Überblick	Velauthapillai, Jeyakrishna
491	2015	SEM modeling with singular moment matrices Part III: GLS estimation	Singer, Hermann