Tilburg University

An alternative derivation of the k-class estimators

Neeleman, D.

Publication date:
1970

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):
Neeleman, D. (1970). *An alternative derivation of the k-class estimators*. (EIT Research Memorandum). Stichting Economisch Instituut Tilburg.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

CBM

76R6

1976/1

7626

1970

21

EIT

21

Bestemming	TIJDSCHRIFTENBUREAU BIBLIOTHEEK KATHOLIEKE HOGESCHOOL TILBURG	Nr.

D. Neeleman

An alternative derivation of the k-class estimators

Research memorandum

R41

T estimation

TILBURG INSTITUTE OF ECONOMICS
DEPARTMENT OF ECONOMETRICS

An Alternative Derivation of the k-Class Estimators.

1. Introduction

In many books on econometrics, for instance, Christ 1966, Dhrymes 1970, Goldberger 1964 and Johnston 1963 the k-Class estimators are introduced with the aid of a system of equations. Only Malinvaud 1966 gives a derivation of this equation system.

In our study another derivation of this system is given which differs from the derivation of Malinvaud on two points.

1. Introducing the Kronecker product of matrices it is possible to transform the simultaneous equation model into a well-known model viz. the general linear model with restrictions.
2. Using the theorems on extrema of quadratic forms it is possible to give a solution of the problem without introducing a normalization rule.

Afterwards one can choose an appropriate normalization rule.

2. Some Theorems on the Extrema of Quadrate Forms

These theorems, on which the derivations in § 4 are based, are given here without proof. The proofs can be found in Rao 1965.

Theorem 1

Let x be an n -vector of real elements.

x_1 be an m -vector of the first m ($< n$) elements of x .

x_2 be an $(n-m)$ -vector of the remaining elements of x .

B be an $n \times k$ matrix of real elements

B_1 be an $m \times k$ matrix of the first m rows of B

B_2 be an $(n-m) \times k$ matrix of the remaining rows of B

c be a k vector of real elements

$\begin{pmatrix} D_1 & D_2 \\ D_3 & D_4 \end{pmatrix}$ be any generalised inverse of $\begin{pmatrix} B_1' & B_1 & B_2' \\ B_2 & 0 \end{pmatrix}$

Then

R 41
(ECO)

$$\inf_{B'x=c} x_1' x_1 = c' D_1 c$$

and the infimum is attained at $x_1^* = B_1 D_1' c$, $x_2^* = D_2' c$ provided that, the linear restrictions $B'x = c$ are consistent.

Theorem 2

Let x be an n -vector of real elements

A be an $n \times n$ symmetric matrix of real elements

B be an $n \times n$ symmetric positive definite matrix of real elements

$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ the roots of $|A - \lambda B| = 0$

Then

$$\inf_x \frac{x'Ax}{x'Bx} = \lambda_1$$

Theorem 3

Let x be an n -vector of real elements

y be a k vector of real elements

A be an $m \times n$ matrix of real elements, $m \geq n$, $\text{rank } A = n$.

B be an $m \times k$ vector of real elements

Then

$$\inf_x 2 x'A'B y + x'A'A x = - y'B'A(A'A)^{-1} A'B y$$

and this infimum is attained at $x^* = - (A'A)^{-1} A'B y$

3. The Model

For the sake of completeness the hypotheses by which the model is described, are briefly formulated below

Hypothesis 1

The model is linear and contains m equations and n variables viz. m endogenous variables $y_1, y_2, \dots, y_m$ and k exogenous variables $z_1, z_2, \dots, z_k$ ($m + k = n$).

Further the model contains m error terms $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_m$ and can be described as

$$By_t + Cz_t = \epsilon_t \quad t = 1, 2, \dots, T$$

$$y_t = \begin{pmatrix} y_{1t} \\ \vdots \\ y_{mt} \end{pmatrix} \quad z_t = \begin{pmatrix} z_{1t} \\ \vdots \\ z_{kt} \end{pmatrix} \quad \epsilon_t = \begin{pmatrix} \epsilon_{1t} \\ \vdots \\ \epsilon_{mt} \end{pmatrix}$$

$$B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mm} \end{pmatrix} \quad C = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1k} \\ c_{21} & c_{22} & \dots & c_{2k} \\ \vdots & \vdots & & \vdots \\ c_{m1} & c_{m2} & \dots & c_{mk} \end{pmatrix}$$

Hypothesis 2

The model is complete i.e. B^{-1} exists

Hypothesis 3

The ϵ_t ($t = 1, 2, \dots, T$) are independent identical distributed random variables with expectation 0 and non-singular covariance matrix Σ .

Hypothesis 4

The equation to be estimated is identified.

Hypothesis 5

The exogenous variables are uniformly bounded i.e. there exists a number K such that for every i ($i = 1, 2, \dots, k$) and every t ($t = 1, 2, \dots, T$) $|z_{it}| < K$ holds

Hypothesis 6

The vector of exogenous variables

$(z_{11}, z_{12}, \dots, z_{1T}), \dots, (z_{k1}, z_{k2}, \dots, z_{kT})$
are linearly independent

This assumption is equivalent to the non-singularity of the matrix $Z'Z$ with $Z' = (z_1, z_2, \dots, z_T)$

Hypothesis 7

$\text{plim}_{T \rightarrow \infty} \frac{1}{T} Z'Z$ exists and is non-singular.

4. The k Class Estimators

The reduced form of the model is

$$y_t = \Pi z_t + \eta_t \quad (t = 1, 2, \dots, T) \quad (1)$$

with

$$\Pi = -B^{-1}C \quad \eta_t = B^{-1} \epsilon_t \quad (2)$$

Starting from the hypotheses 1,2 and 3 given in section 3 one can derive the following properties for the error terms of the reduced form η_t :

The η_t ($t = 1, 2, \dots, T$) are independent identical distributed random variables with expectation 0 and non-singular covariance matrix $(B^{-1}) \Sigma (B^{-1})' = \Omega$.

Suppose that one wishes to estimate the first equation of the structural form of the model. Without limiting the validity of the conclusions one can state that the number of endogenous variables in this equation equals m .

Because of the order condition this structural equation is identifiable only if the number of exogenous variables excluded from this equation is at least as great as $(m-1)$ i.e. in the first row of C at least $(m-1)$ zero's appear.

Rearranging the variables such that these zero's occupy the last places in the first row of C , the first equation of the structural form can be written as

$$b'x_t + (c', 0) z_t = \epsilon_t \quad (t = 1, 2, \dots, T) \quad (3)$$

where

b' is the first row of B

$(c', 0)$ is the first row of C

The ordinary least squares estimators of b and c are not unbiased and even not consistent hence one has to search for another estimation method. For that purpose the model is written, using the Kronecker product of matrices, as

$$\text{vec } Y = (I_m \quad Z) \text{vec } \Pi' + \text{vec } \eta \quad (4)$$

where

$$Y = \begin{pmatrix} y'_1 \\ \vdots \\ y'_T \end{pmatrix} \quad Z = \begin{pmatrix} z'_1 \\ \vdots \\ z'_T \end{pmatrix} \quad \eta = \begin{pmatrix} \eta'_1 \\ \vdots \\ \eta'_T \end{pmatrix} \quad (5)$$

The restriction (3) can be written like

$$(b' \otimes \frac{T}{k}) \text{vec } \Pi' = \begin{pmatrix} -c \\ 0 \end{pmatrix} \quad (6)$$

For the properties of the Kronecker product we refer to Neudecker 1968.

The problem can be solved as follows

Step 1

Estimate, with the aid of generalised least squares $\text{vec } \Pi'$ starting from known b and c taking into account the restriction (6)

In other words minimize

$$S = [\text{vec } Y - (I_m \otimes Z) \text{vec } \Pi']' (\Omega^{-1} \otimes I_T) [\text{vec } Y - (I_m \otimes Z) \text{vec } \Pi'] \quad (7)$$

subject to the restriction in (6)

Step 2

Calculate the minimum of S in step 1. This minimum is a function of b and c . Next calculate those values of b and c that realise the minimum of this function

(i) The problem as it is formulated in step 1 can be solved with the aid of theorem 1 in section 2 by formulating it as follows:

Determine the linear function

$$q' \text{ vec } Y + d$$

so that

$$V(q' \text{ vec } Y + d) = q' (\Omega \otimes I_T) q \quad (8)$$

is minimum under the constraints

$$E(q' \text{ vec } Y + d) = p' \text{ vec } \Pi' \quad (9)$$

$$(b' \otimes I_k) \text{ vec } \Pi' = \begin{pmatrix} -c \\ 0 \end{pmatrix} \quad (10)$$

where p is an arbitrary vector.

The condition (9) implies that

$$q' (I_m \otimes Z) \text{ vec } \Pi' + d = p' \text{ vec } \Pi' \quad (11)$$

Then from (10) and (11) it follows that there exists a vector r such that

$$q' (I_m \otimes Z) + r' (b' \otimes I_k) = p' \quad (12)$$

and

$$d = r' \begin{pmatrix} -c \\ 0 \end{pmatrix} \quad (13)$$

The problem then reduces to minimizing

$$q'(\Omega \otimes I_T) q \quad (14)$$

subject to the constraint

$$(I_m \otimes Z', b \otimes I_k) \begin{pmatrix} q \\ r \end{pmatrix} = p \quad (15)$$

Or taking into account that

$$(\Omega \otimes I_T) = (WW' \otimes I_T) \quad (16)$$

where W is a nonsingular $m \times m$ matrix and taking

$$(W' \otimes I_T) q = s \quad (17)$$

the problem is transformed into minimizing

$$s's \quad (18)$$

subject to the constraint

$$[(W^{-1})' \otimes Z', b \otimes I_k] \begin{pmatrix} s \\ r \end{pmatrix} = p \quad (19)$$

Application of theorem 1 in section 2 gives

$$s^* = (W^{-1} \otimes Z) D_1' p \quad \text{or} \quad q^* = (\Omega^{-1} \otimes Z) D_1' p \quad (20)$$

$$r^* = D_2' p$$

where D_1 and D_2 are submatrices of the generalised inverse

$$\begin{pmatrix} \Omega^{-1} \otimes Z'Z & b \otimes I_k \\ b' \otimes I_k & 0 \end{pmatrix}^{-1} = \begin{pmatrix} D_1 & D_2 \\ D_3 & D_4 \end{pmatrix} \quad (21)$$

The minimum variance unbiased estimator of $p' \text{vec } \Pi'$ is

$$p' \text{vec } \hat{\Pi}' = p' D_1 (\Omega^{-1} \otimes Z) \text{vec } Y + p' D_2 \begin{pmatrix} -c \\ 0 \end{pmatrix} \quad (22)$$

where $\text{vec } \hat{\Pi}'$ is the solution of the equations

$$\begin{aligned} (\Omega^{-1} \otimes Z'Z) \text{vec } \Pi' - (b \otimes I_k) \lambda &= (\Omega^{-1} \otimes Z) \text{vec } Y \\ (b' \otimes I_k) \text{vec } \Pi' &= \begin{pmatrix} -c \\ 0 \end{pmatrix} \end{aligned} \quad (23)$$

Solution of this equations gives

$$\begin{aligned} \text{vec } \hat{\Pi}' &= [I_m \otimes (Z'Z)^{-1} Z'] \text{vec } Y - [(b' \Omega b)^{-1} \Omega b b' \otimes (Z'Z)^{-1} Z'] \text{vec } Y \\ &\quad - [(b' \Omega b)^{-1} \Omega b \otimes I_k] \begin{pmatrix} c \\ 0 \end{pmatrix} \end{aligned} \quad (24)$$

(ii) In step 2 we substitute $\text{vec } \hat{\Pi}'$ into (7) which yields

$$\begin{aligned} S &= (\text{vec } Y)' (\Omega^{-1} \otimes Z') \text{vec } Y - (\text{vec } Y)' [\Omega^{-1} \otimes Z(Z'Z)^{-1} Z'] \text{vec } Y \\ &\quad + (\text{vec } Y)' [(b' \Omega b)^{-1} b b' \otimes Z(Z'Z)^{-1} Z'] \text{vec } Y + \quad (25) \\ &\quad + 2(c', 0) [(b' \Omega b)^{-1} b' \otimes Z'] \text{vec } Y + (c', 0) [(b' \Omega b)^{-1} \otimes Z'Z] \begin{pmatrix} c \\ 0 \end{pmatrix} \end{aligned}$$

b and c has to be chosen in such way that S becomes minimum. This means that one can only consider that part of S in which b and c appear.

This part can be written as

$$S_{\text{part}} = \frac{b' Y' Z (Z'Z)^{-1} Z' Y b + 2c' Z_1' Y b + c' Z_1' Z_1 c}{b' \Omega b} \quad (26)$$

where Z_1 is the matrix of the first k_1 columns of Z (k_1 is the number of exogenous variables belonging to the first equation).

Now it is easy to see that determining the minimum of S_{part} depends only on the choice of b . For from theorem 3 section 2 follows that

$$\inf_c 2 c' Z_1' Y b + c' Z_1' Z_1 c = - b' Y' Z_1 (Z_1' Z_1)^{-1} Z_1' Y b \quad (27)$$

and this infimum is attained at

$$\hat{c} = - (Z_1' Z_1)^{-1} Z_1' Y b \quad (28)$$

So the problem reduces to

Determine

$$\inf_b \frac{b' [Y' Z (Z' Z)^{-1} Z' Y - Y' Z_1 (Z_1' Z_1)^{-1} Z_1' Y] b}{b' \Omega b} \quad (29)$$

From theorem 2 in section 2 it follows immediately that the infimum of (29) equals λ_1 , where λ_1 is the smallest root of

$$|Y' Z (Z' Z)^{-1} Z' Y - Y' Z_1 (Z_1' Z_1)^{-1} Z_1' Y - \lambda \Omega| = 0 \quad (30)$$

This means that for every $\hat{b}$ that is a solution of

$$[Y' Z (Z' Z)^{-1} Z' Y - Y' Z_1 (Z_1' Z_1)^{-1} Z_1' Y - \lambda_1 \Omega] \hat{b} = 0 \quad (31)$$

the infimum is attained.

The equations (28) and (31) can be combined to

$$\begin{pmatrix} Y' Z (Z' Z)^{-1} Z' Y - \lambda_1 \Omega_1 & Y' Z_1 \\ Z_1' Y & Z_1' Z_1 \end{pmatrix} \begin{pmatrix} \hat{b} \\ \hat{c} \end{pmatrix} = 0 \quad (32)$$

If we take as normalization rule that the first element of $\hat{b}$ equals -1 then the system of equations (32) changes into

$$\begin{pmatrix} Y_2' Z (Z' Z)^{-1} Z' Y_2 - \lambda_1 \Omega_{22} & Y_2' Z_1 \\ Z_1' Y_2 & Z_1' Z_1 \end{pmatrix} \begin{pmatrix} \bar{b}_2 \\ \bar{c} \end{pmatrix} = \begin{pmatrix} Y_2' Z (Z' Z)^{-1} Z' Y_1 - \lambda \Omega_{21} \\ Z_1' Y_1 \end{pmatrix} \quad (33)$$

where $\begin{pmatrix} \bar{b}_2 \\ \bar{c} \end{pmatrix}$ are the remaining unknown elements of $\begin{pmatrix} \bar{b} \\ \bar{c} \end{pmatrix}$
after normalization

Y_1 is the first row of Y

Y_2 is the matrix of the remaining rows of Y

Ω_{21} is the first column of Ω minus the first element

Ω_{22} is the matrix of the last $(m-1)$ rows and the last $(m-1)$
columns

If Ω is known λ_1 , $\bar{b}_2$ and $\bar{c}$ can be calculated from (30) and (33). In general however Ω will be unknown but can be consistently estimated by

$$\hat{\Omega} = \frac{V'V}{T} \quad (34)$$

where

$$V = Y - Z(Z'Z)^{-1} Z'Y \quad (35)$$

In this case (30) becomes

$$\left| Y'Z(Z'Z)^{-1}Z'Y - Y'Z_1(Z_1'Z_1)^{-1}Z_1'Y - \lambda \frac{V'V}{T} \right| = 0 \quad (36)$$

and (33) becomes

$$\begin{pmatrix} Y_2'Y_2 - (1 + \frac{\lambda_1}{T})V_2'V_2 & Y_2'Z_1 \\ Z_1'Y_2 & Z_1'Z_1 \end{pmatrix} \begin{pmatrix} \bar{b}_2 \\ \bar{c} \end{pmatrix} = \begin{pmatrix} Y_2'Y_1 - (1 + \frac{\lambda_1}{T})V_2'V_1 \\ Z_1'Y_1 \end{pmatrix} \quad (37)$$

where $V_2'V_2$ en $V_2'V_1$ are submatrices of $V'V$ that correspond with Ω_{22}
en Ω_{21}

It can be proved, see for instance Goldberger 1964 that $\text{plim}_{T \rightarrow \infty} (1 + \frac{\lambda_1}{T}) = 1$
and that $\bar{b}_2$ and $\bar{c}$ are consistent estimators.

Replacing $(1 + \frac{\lambda_1}{T})$ by k one finds

$$\begin{pmatrix} Y_2'Y_2 - kV_2'V_2 & Y_2'Z_1 \\ Z_1'Y_2 & Z_1'Z_1 \end{pmatrix} \begin{pmatrix} \bar{b}_2 \\ \bar{c} \end{pmatrix} = \begin{pmatrix} Y_2'Y - kV_2'V_1 \\ Z_1'Y_1 \end{pmatrix} \quad (38)$$

which are the equations of the so-called k class estimators. It is easy to see that if $\text{plim}_{T \rightarrow \infty} k = 1$ these estimators are consistent.

4. References

- Christ C.F. (1966) *Econometric Models and Methods*, New York Wiley 1966
- Dhrymes P.J. 1970 *Econometrics*, New York, Harper and Row 1970
- Goldberger A.S. (1964), *Econometric Theory*, New York: Wiley 1964
- Johnston J. (1963), *Econometric Methods* New York: McGraw-Hill 1963
- Malinvaud E. (1966), *Statistical Methods of Econometrics*, Amsterdam: North Holland Publishing Company 1966
- Nendecker H. (1968), The Kronecker Matrix Product and Some of its Applications in Econometrics
Statistica Neerlandica, Vol. 22 no. 1 1968
- Rao C.R. (1965) *Linear Statistical Inference and Its Applications*.
 New York: Wiley 1965

PREVIOUS NUMBERS:

EIT 1	J. Kriens *)	Het verdelen van steekproeven over subpopulaties bij accountantscontroles.
EIT 2	J. P. C. Kleynen *)	Een toepassing van „Importance sampling”.
EIT 3	S. R. Chowdhury and W. Vandaele *)	A bayesian analysis of heteroscedasticity in regression models.
EIT 4	Prof. drs. J. Kriens *)	De besliskunde en haar toepassingen.
EIT 5	Prof. dr. C. F. Scheffer	Winstkapitalisatie versus dividendkapitalisatie bij het waarderen van aandelen.
EIT 6	S. R. Chowdhury *)	A bayesian approach in multiple regression analysis with inequality constraints.
EIT 7	P. A. Verheyen	Investeren en onzekerheid.
EIT 8	R. M. J. Heuts en Walter H. Vandaele	Problemen rond niet-lineaire regressie.
EIT 9	S. R. Chowdhury	Bayesian analysis in linear regression with different priors.
EIT 10	A. J. van Reeken	The effect of truncation in statistical computation.
EIT 11	W. H. Vandaele and S. R. Chowdhury	A revised method of scoring.
EIT 12	J. de Blok	Reclame-uitgaven in Nederland.
EIT 13	Walter H. Vandaele	Mødsco, a computer program for the revised method of scoring.
EIT 14	J. Plasmans	Alternative production models. (Some empirical relevance for postwar Belgian Economy)
EIT 15	D. Neeleman	Multiple regression and serially correlated errors.
EIT 16	H. N. Weddepohl	Vector representation of majority voting.
EIT 17	Walter H. Vandaele	Zellner's seemingly unrelated regression equation estimators: a survey.
EIT 18	J. Plasmans	The general linear seemingly unrelated regression problem. I. Models and Inference.
EIT 19	J. Plasmans and R. Van Straelen .	The general linear seemingly unrelated regression problem. II. Feasible statistical estimation and an application.
EIT 20	Pieter H. M. Ruys	A procedure for an economy with collective goods only.

*) not available

EIT 1970