45a

PSYCHOMETRIKA—VOL. 53, NO. 3, 407–415
SEPTEMBER 1988
COMPUTATIONAL PSYCHOMETRICS

FACAIC: MODEL SELECTION ALGORITHM FOR THE ORTHOGONAL FACTOR MODEL USING AIC AND CAIC

HAMPARSUM BOZDOGAN AND DONALD E. RAMIREZ

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF VIRGINIA

This paper describes the authors' FORTRAN algorithm FACAIC for choosing the number of factors for an orthogonal factor model using Akaike's Information Criterion. FACAIC utilizes the IMSL subroutine OFCOMM.

Key words: factor analysis algorithm, choosing the number of factors, AIC and CAIC.

Purpose and Description

Purpose

Factor analysis is a very useful and important multivariate technique which is used to find a way of condensing or summarizing the information contained in a number of original variables into a smaller set of composite dimensions, factors, or hypothetical constructs with minimum loss of information.

In the statistical literature, an exact satisfactory quantitative technique for deciding how many factors to extract is largely unresolved, and does not seem to be available, except some of the ad hoc stopping rules which are currently being utilized. The classical theory of hypothesis testing is not appropriate in the context of factor model problems, since the problem is not of testing a particular hypothesis, but is rather a *multiple decision problem*. The usual likelihood ratio test is not always valid, and the significance level for the test criterion is not adjusted in the sequential testing process.

Therefore, the algorithm presented in this paper uses Akaike's (1974) information criterion (AIC) and the asymptotically consistent Akaike's information criterion (CAIC) to choose the number of factors. These being *information-theoretic* identification procedures, the researcher is forced to be more neutral in the choice of the model and the naming of the common factors.

Description

The factor analysis model assumes a structure of the form:

$$\underset{(p\times 1)}{\mathbf{X}} = \underset{(p\times m)}{\Lambda}\ \underset{(m\times 1)}{\mathbf{f}} + \underset{(p\times 1)}{\boldsymbol{\mu}} + \underset{(p\times 1)}{\boldsymbol{\varepsilon}} \tag{1}$$

where

Λ = the factor loading matrix, that is, a matrix of unknown coefficients λ_{ij},

$\mathbf{f}$ = a random vector of common factors,

The authors dedicate this algorithm to Professor Hirotugu Akaike in appreciation of his pioneering work on AIC which was originally intended for the factor analysis and other statistical model identification problems.

Requests for reprints should be sent to Hamparsum Bozdogan, Department of Mathematics, University of Virginia, Charlottesville, VA 22903.

0033-3123/88/0900-a011$00.75/0

$\boldsymbol{\mu}$ = mean vector of p variables,

$\boldsymbol{\varepsilon}$ = a random vector of *specific* or *unique* factors, or the error term,

and with

$\mathbf{f}$ and $\boldsymbol{\varepsilon}$ independent

$E(\mathbf{f}) = 0,$

$\text{Cov}(\mathbf{f}) = \mathbf{I},$

$E(\boldsymbol{\varepsilon}) = 0,$

$\text{Cov}(\boldsymbol{\varepsilon}) = \psi,$ where ψ is a diagonal matrix.

The algorithm is presented as a FORTRAN 77 subroutine. The user is assumed to have computed the correlation matrix R from the data matrix X. The user inputs to the subroutine FACAIC are:

N : INTEGER
NUMBER OF CASES

P : INTEGER
NUMBER OF VARIABLES

DOAIC : LOGICAL
.TRUE. FOR AIC
.FALSE. FOR CAIC

CORR : DOUBLE PRECISION VECTOR OF DIMENSION $(P*(P + 1)/2)$
CORRELATION MATRIX IN SYMMETRIC STORAGE

AUTO : LOGICAL
.TRUE. FOR AUTOMATIC SEARCH FOR BEST MODEL
.FALSE. FOR MODEL PREDETERMINED BY USER

SETNF : LOGICAL
NUMBER OF FACTORS PREDETERMINED BY USER
WHEN AUTO = .FALSE.

The outputs of the subroutine FACAIC are:

BESTNF : INTEGER
BEST NUMBER OF FACTORS

ALPHA : DOUBLE PRECISION
P-VALUE FOR GOODNESS-OF-FIT

NPARMS : DOUBLE PRECISION
NUMBER OF PARAMETERS

PENALT : DOUBLE PRECISION
PENALTY COMPONENT FOR AIC OR CAIC USING
BESTNF FACTORS

AIC : DOUBLE PRECISION
AKAIKE'S INFORMATION CRITERION FOR MODEL
SELECTION

CAIC :

LNDET :

PSI :

CRITER :

LAMBDA :

INVCOR :

COMM :

RESCOR :

IEROFC :

LNLIKE :

NITERS :

FMIN :

CHISQ :

Since the ma
correlation matrix

on the original ob
be between 1 and

CAIC : DOUBLE PRECISION
CONSISTENT AIC FOR MODEL SELECTION

LNDET : DOUBLE PRECISION
NATURAL LOG OF DETERMINANT OF CORRELATION MATRIX

PSI : DOUBLE PRECISION VECTOR OF DIMENSION (*P*)
PSI VECTOR OF ERROR VARIANCES

CRITER : DOUBLE PRECISION MATRIX OF DIMENSION (0:*P*,5)
MATRIX OF NUMBER OF PARAMETERS, AIC VALUES, CAIC VALUES,
DEGREES-OF-FREEDOM, AND *P*-VALUES FOR EACH MODEL.

LAMBDA : DOUBLE PRECISION MATRIX OF DIMENSION (*P*,*P*)
LAMBDA MATRIX OF FACTOR LOADINGS

INVCOR : DOUBLE PRECISION VECTOR OF DIMENSION (*P**(*P* + 1)/2)
INVERSE OF CORRELATION MATRIX IN SYMMETRIC STORAGE

COMM : DOUBLE PRECISION VECTOR OF DIMENSION (*P*)
VECTOR OF COMMUNALITIES

RESCOR : DOUBLE PRECISION VECTOR OF DIMENSION (*P**(*P* + 1)/2)
NORMAL RESIDUAL CORRELATION MATRIX IN SYMMETRIC STORAGE

IEROFC : INTEGER VECTOR OF DIMENSION (0:*P*)
IMSL ERROR FLAG FOR OFCOMM FOR EACH MODEL

LNLIKE : DOUBLE PRECISION
SCALED LOG-LIKELIHOOD

NITERS : INTEGER
NUMBER OF ITERATIONS USED IN OFCOMM IN FINAL MODEL

FMIN : INTEGER
VALUE OF THE FUNCTION MINIMUM UNDER THE MAXIMUM LIKELIHOOD PROCEDURE
IN OFCOMM IN FINAL MODEL

CHISQ : DOUBLE PRECISION
THE CHI-SQUARED STATISTIC IN OFCOMM IN FINAL MODEL

Since the maximum likelihood factor model is scale invariant, the program uses the correlation matrix R in estimation of Λ and Ψ, and in extracting the factors such that

$$R \doteq \hat{\Lambda}\hat{\Lambda}' + \hat{\Psi}, \tag{2}$$

on the original observable random vector **X**. The number of factors NF are constrained to be between 1 and BIGFAC = .5 * (2 * p + 1 − (8 * p + 1) ** .5). As the models increase in size, the Ψ (PSI) values from the previous model are used as starting values for the next larger model to achieve much faster convergence, to obtain more stable and efficient numerical solutions in the iteration process. To make Λ (LAMBDA) well defined a computationally convenient *uniqueness* condition is imposed on Λ to force

$$\Lambda'\Psi^{-1}\Lambda = \Delta, \tag{3}$$

where Δ is a diagonal matrix. For more details, see, for example, Johnson and Wichern (1988).

Given the number of factors NF, minus twice the log of the maximized likelihood function at the convergence of the two-step Newton-Raphson iterative procedure, is given by

$$-2\,\ell n\, L(\hat{\mu}, \hat{\Lambda}, \hat{\Psi}) = N * P * \ell n\,(2 * \pi) + N * \,\ell n\,(\det R) + N * P. \tag{4}$$

As NF increases, minus twice the log of the maximized likelihood function in (4) increases. This expression in (4), constitutes the first component of the AIC model selection criterion, which is a measure of *badness of fit*, *inaccuracy*, or *bias* when the maximum likelihood estimators of the parameters of the model are used. To compensate the bias in (4), or to adjust the increased unreliability, AIC has a second component which is called the *penalty* component or the *measure of complexity*, which is usually 2∗ (number of independent parameters within the model). Thus

$$\begin{aligned}\mathrm{AIC}(NF) = {} & N * P * \ell n\,(2 * \pi) + N * \ell n\,(\det R) + N * P \\ & + 2[(NF * P + P) - .5 * NF(NF - 1)].\end{aligned} \tag{5}$$

When $NF = 0$, for us the factor model is *unrestricted*, and NPARMS, the number of parameters is $P(P + 1)/2$.

The AIC statistic is a new procedure that is free from the ambiguities inherent in the application of conventional hypothesis testing procedures. AIC balances the risk due to the bias when a lower dimension is selected and the risk due to the increase of variance when a higher dimension is selected.

In the literature, other penalties have been proposed by Schwarz (1978), Akaike (1978), Bozdogan (1985), and by other researchers to make AIC asymptotically consistent to penalize over-parameterization more stringently to pick only the simplest of the true models whenever there is nothing to be lost by doing so. Here the penalty is the number of parameters times $\ell n\,(N)$, the natural logarithm of the sample size N. In this case, the model selection criterion is called CAIC, a generic name dubbed by Bozdogan, which is a consistent version of AIC, and one of its forms is defined by

$$\mathrm{CAIC}(NF) = N * P * \ell n\,(2 * \pi) + N * \ell n\,(\det R) + N * P + \mathrm{NPARMS} * \ell n\,(N). \tag{6}$$

Thus, the CAIC criterion favors lower dimensional models and achieves a fully automatic Occam's Razor, that is, choosing the simplest of the true models, whatever they might be. The AIC criterion also chooses the parsimonious models, but it is not as stringent as the CAIC. In this sense we use AIC to determine the upper bound and use CAIC to determine the lower bound of the number of factors to be selected which gives an interval or range on choosing the number of factors. For each model, both criteria are computed along with the goodness-of-fit for the model with NF factors.

After the program has found the correct number of factors to be used, the correlation structure is rotated according to the VARIMAX criterion of Kaiser (1958). This *spreads out* the squares of the loadings on each factor as much as possible. The entries in LAMBDA estimate the correlations between the variables and the common factors.

TABLE 1. *Correlations Amo*

Variable
1. Height
2. Arm span
3. Length of forearm
4. Length of lower leg
5. Weight
6. Bitrochanteric diameter
7. Chest girth
8. Chest width

For more details on the mo
refer the reader to Bozdogan and

Most of the computations t
Inc., Houston, TX). A brief descr

Subroutine	
LUDECP	For c
OFCOMM	For c
OFROTA	For c
UERSET	For s

The convergence criteria for
IMSL.

TABLE 2. *The AIC's, CAIC's*
Physical Variable

No. of Factors $m = NF$	No. of Para $s = NP$
1	16
2	23
3	29
4	34
5	38

NOTE: $n = 305$, $p = 8$

*Minimum AIC at $\hat{m} = 4$ fa

**Minimum CAIC at $\hat{m} = 3$ f

***P-Value of χ^2 cannot be

AIC (Unrestricted Model)

CAIC (Unrestricted Model

TABLE 1. *Correlations Among Eight Physical Variables for 305 Girls*

Variable	1	2	3	4	5	6	7	8
1. Height	-	-	-	-	-	-	-	-
2. Arm span	.846	-	-	-	-	-	-	-
3. Length of forearm	.805	.881	-	-	-	-	-	-
4. Length of lower leg	.859	.826	.801	-	-	-	-	-
5. Weight	.473	.376	.380	.436	-	-	-	-
6. Bitrochanteric diameter	.398	.326	.319	.329	.762	-	-	-
7. Chest girth	.301	.277	.237	.327	.730	.583	-	-
8. Chest width	.382	.415	.345	.365	.629	.577	.539	-

For more details on the model selection approach to the factor model problem, we refer the reader to Bozdogan and Ramirez (1986, 1987).

Related Algorithms

Most of the computations take place in subroutines from the IMSL package (IMSL, Inc., Houston, TX). A brief description of the purpose of each is given below.

Subroutine	*Purpose*
LUDECP	For computing determinants of symmetric matrices
OFCOMM	For computing unrotated factor loading matrix
OFROTA	For orthogonal rotation of a factor loading matrix
UERSET	For suppression of warning errors

The convergence criteria for OFCOMM and OFROTA are those recommended by IMSL.

TABLE 2. *The AIC's, CAIC's, and the P-Values for χ^2 for the Eight Physical Variables Data Under the Orthogonal Factor Model*

No. of Factors m = NF	No. of Parameters s = NP	AIC(m)	CAIC(m)	DF	P-Values of χ^2
1	16	5461.4	5520.9	20	0.00000
2	23	4930.7	5016.2	13	0.00000
3	29	4888.5	4996.4**	7	0.00199
4	34	4879.9*	5006.4	2	0.11538
5	38	4883.5	5024.9	-2	***

NOTE: $n = 305$, $p = 8$

*Minimum AIC at $\hat{m} = 4$ factors.

**Minimum CAIC at $\hat{m} = 3$ factors.

***P-Value of χ^2 cannot be computed due to nonpositive d.f.

AIC (Unrestricted Model) = 4879.4

CAIC (Unrestricted Model) = 5013.4

TABLE 3. *Maximum Likelihood Solutions for the Eight Physical Variables Data (Four Common Factors)*

Variable j	f_1	f_2	f_3	f_4	$h_j^2 = 1-\hat{\Psi}_j$
1	0.87767	0.28183	-0.11340	-0.00831	0.86266
2	0.93719	0.19848	0.28309	0.04619	1.00000
3	0.87324	0.19563	0.08512	-0.00687	0.80811
4	0.88602	0.21437	-0.18783	0.13540	0.88459
5	0.24143	0.88324	-0.10771	0.10642	0.86134
6	0.18241	0.82316	-0.01638	-0.08007	0.71754
7	0.11340	0.73252	0.00473	0.51949	0.81934
8	0.25972	0.64603	0.14001	0.08089	0.51096

A Numerical Example

To show the practicality and versatility of the FACAIC algorithm we shall give a numerical example by using the well-known empirical data set of Mullen (1939) which consist of $p = 8$ physical variables on $n = 305$ girls from seven to seventeen years of age. For this, we also refer the reader to Harman (1976, p. 22).

The correlation matrix for this data set is given in Table 1.

Using FACAIC, we performed a maximum likelihood factor analysis by fitting $m = 1, 2, 3, 4$, and 5 factors. Our results are given in Table 2 along with the results for the unrestricted model.

Looking at Table 2, we see that the minimum AIC occurs at $m = 0$, indicating that the unrestricted model is consistent and parsimonious with the data. Since $m = 4$ has nearly the same value, we treat $m = 4$ factor model as also consistent and parsimonious with the data. On the other hand, the minimum CAIC gives $m = 3$ factors as the best fitting factor model. But we note that the chi-squared value for $m = 3$ is 22.6 with $DF = 7$ and P-Value = .00199, which is not significant as compared with the P-Value of $m = 4$ factor model. Therefore, for now tentatively considering that $m =$ 4-factor model is the

TABLE 4. *Two Maximum Likelihood Solutions for the Eight Physical Variables Data (Three Common Factors)*

Variable	Harman Solution				Jennrich & Robinson Solution			
j	f_1	f_2	f_3	h_j^2	f_1	f_2	f_3	h_j^2
1	.862	.320	.164	.873	.846	.189	-.348	.873
2	.867	.432	-.243	.998	1.000	.000	.000	1.000
3	.808	.388	-.051	.807	.881	.055	-.164	.806
4	.833	.343	.180	.843	.826	.160	-.368	.843
5	.748	-.584	.090	.910	.376	.876	.031	.910
6	.625	-.499	.007	.640	.326	.725	.093	.641
7	.569	-.515	-.020	.589	.277	.704	.130	.589
8	.603	-.344	-.164	.509	.415	.543	.204	.509

TABLE 5. *Maximum Likelihood Solutions for the Eight Physical Variables Data (Three Common Factors)*

Variable	Bozdogan and Ramirez Solution			
j	f_1	f_2	f_3	h_j^2
1	0.883	0.275	-0.129	0.873
2	0.936	0.204	0.287	1.000
3	0.872	0.196	0.088	0.806
4	0.875	0.239	-0.145	0.844
5	0.234	0.919	-0.105	0.910
6	0.185	0.779	-0.021	0.641
7	0.130	0.757	0.004	0.589
8	0.254	0.648	0.157	0.509

parsimonious one, in Table 3 we give our solutions for $m = 4$-factor model which contains the columns of the estimated factor loading matrix $\hat{\Lambda}$ and the communalities $h_j^2 = 1 - \hat{\Psi}_j$.

Looking at Table 3, we see that Variable 2 has communality = 1.00 (or equivalently $\hat{\Psi}_2 = 0.00$), indicating that there seems to be a *Heywood case*, (Heywood, 1931), that is, an *improper solution* when we fit $m = 4$-factor model.

In such a case, when we fit a reduced number of factors, say, for example, $m = 3$-factor model, in Table 4 we show the results of Harman (1976, p. 212) and Jennrich and Robinson (1969), and in Table 5, we show our results for fitting $m = 3$-factor model.

Comparing the results in Table 4 and in Table 5, for three common factors, we note that all tree solutions have nearly identical communalities and thus have the same Ψ (PSI) matrix. Still Variable 2 has communality = 1, indicating the presence of a Heywood case even in $m = 3$-factor model. Hence, we suggest a reanalysis of the data this time with Variable 2 deleted from the data matrix. In doing so, our new results on AIC and CAIC values are given in Table 6, and the maximum likelihood solutions are given in Table 7.

TABLE 6. *The AIC's, CAIC's, and the P-Values for χ^2 for the Seven Physical Variables Data (Variable 2 Removed)*

m = NF	s = NP	AIC(m)	CAIC(m)	DF	P-Values of χ^2
1	14	5084.2	5136.3	14	0.00000
2	20	4580.5	4654.9**	8	0.00727
3	25	4570.7*	4663.7	3	0.67651
4	29	4577.1	4685.0	-1	***

NOTE: $n = 305$, $p = 7$

*Minimum AIC at $\hat{m} = 3$

**Minimum CAIC at $\hat{m} = 2$

***P-Value of χ^2 cannot be computed due to nonpositive d.f.

AIC (Unrestricted Model) = 4575.1

CAIC (Unrestricted Model) = 4679.3

TABLE 7. *Maximum Likelihood Solutions for the Seven Physical Variables Data (Three Common Factors, Variable 2 Removed)*

Variable j	f_1	f_2	f_3	h_j^2
1	0.89089	0.27344	-0.00245	0.86846
3	0.84573	0.18846	-0.00023	0.75077
4	0.90186	0.20417	0.14231	0.87528
5	0.26391	0.86112	0.15594	0.83549
6	0.18795	0.84404	-0.08981	0.75580
7	0.11920	0.71858	0.51164	0.79234
8	0.24417	0.63772	0.10057	0.47642

Table 7 clearly shows that the first factor (f_1) is associated with variable (1) *height*, (3) *length of forearm*, and (4) *length of lower leg*, and is thus called the "lankiness" factor. The second factor (f_2) is associated with (5) *weight*, (6) *bitrochanteric diameter* (i.e., the diameter of trochanteric home which is part of the hip joint), (7) *chest girth*, and (8) *chest width*, and has been named the "stockiness" factor. Our analysis shows that the third factor (f_3) is associated with Variable (7) *chest girth*, and we have named it "chestiness" factor.

In Table 8, we give the model correlations for this solution.

The normalized residual correlation matrix is given in Table 9. These values are very helpful in finding out why a model does not fit a data correlation matrix. The off-diagonal values are defined to be

$$\frac{R_{ij} - \hat{R}_{ij}}{(\Psi_i \Psi_j)^{1/2}}. \tag{7}$$

For example, for Variables 3 and 8, we compute

$$\frac{R_{38} - \hat{R}_{38}}{(\Psi_3 \Psi_8)^{1/2}} = \frac{.345 - .32666}{[(1 - .75077)(1 - .47642)]^{1/2}} = .05077. \tag{8}$$

Thus, in concluding, we see that by using the FACAIC algorithm, we can decide on how many factors it is worth fitting to the data; we can identify very readily the Heywood cases (or *improper solutions*; see, e.g., also Akaike, 1986); and we can determine the appropriate number of factors leading to a *parsimony of description*, providing the re-

TABLE 8. *Model Correlations Among Seven Physical Variables (Variable 2 Removed)*

Variable	1	3	4	5	6	7	8
1. Height	-	-	-	-	-	-	-
3. Length of forearm	.805	-	-	-	-	-	-
4. Length of lower leg	.859	.801	-	-	-	-	-
5. Weight	.470	.385	.436	-	-	-	-
6. Bitrochanteric diameter	.398	.318	.329	.762	-	-	-
7. Chest girth	.301	.236	.327	.730	.583	-	-
8. Chest width	.392	.327	.365	.629	.575	.539	-

TABLE 9. *Normal Variab*

Variable
1. Height
3. Length of for
4. Length of lo
5. Weight
6. Bitrochanteri
7. Chest girth
8. Chest width

searcher an analytic up until now.

Akaike, H. (1974). A new
AC-19, 716–723.
Akaike, H. (1978). A Ba
Mathematics, 30 (Pa
Akaike, H. (1986). Factor
Bozdogan, H. (1985). Mo
analytical extensions
Department of Math
Bozdogan, H., & Ramir
choosing the correct
on Multivariate Stat
Sciences—ASA Virgi
Bozdogan, H., & Ramirez
structure in factor an
and data analysis (pp.
Harman, H. (1976). Moder
Heywood, H. B. (1931).
486–501.
Jennrich, R. I., & Robins
sis. Psychometrika, 34,
Johnson, R., & Wichern, D
Kaiser, M. (1958). The var
Mullen, F. (1939). Facto
dissertation, Departm
Schwarz, G. (1978). Estima

TABLE 9. *Normalized Residual Correlation Matrix for the Seven Physical Variables (Variable 2 Removed)*

Variable	1	3	4	5	6	7	8
1. Height	-	-	-	-	-	-	-
3. Length of forearm	.0001	-	-	-	-	-	-
4. Length of lower leg	.0005	-.0010	-	-	-	-	-
5. Weight	.0191	-.0269	-.0001	-	-	-	-
6. Bitrochanteric diameter	-.0026	.0039	-.0003	-.0021	-	-	-
7. Chest girth	-.0026	.0039	-.0002	-.0001	.0002	-	-
8. Chest width	-.0368	.0508	.0011	-.0009	.0053	.0006	-

searcher an analytical quantitative multiple decision procedure which was not available up until now.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, AC-19*, 716–723.

Akaike, H. (1978). A Bayesian analysis of the minimum AIC procedure. *Annals of the Institute of Statistical Mathematics, 30* (Paper A), 9–14.

Akaike, H. (1986). Factor analysis and AIC. *Psychometrika, 52*, 317–332.

Bozdogan, H. (1985). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions (Technical Paper No. 7 in Statistics). University of Virginia, Charlottesville, VA: Department of Mathematics.

Bozdogan, H., & Ramirez, D. E. (1986, May). *The effect of multivariate nonnormality on the probability of choosing the correct number of factors using AIC and CAIC.* A paper presented at the Advanced Symposium on Multivariate Statistical Modeling and Data Analysis, the 64th Meeting of the Virginia Academy of Sciences—ASA Virginia Chapter at James Madison University, Harrisonburg, Virginia.

Bozdogan, H., & Ramirez, D. E. (1987). An expert model selection approach to determine the "best" pattern structure in factor analysis models. In H. Bozdogan & A. K. Gupta (Eds.), *Multivariate statistical modeling and data analysis* (pp. 35–60). Dordrecht, Holland: D. Reidel.

Harman, H. (1976). *Modern factor analysis*, Chicago, IL: The University of Chicago Press.

Heywood, H. B. (1931). On finite sequences of real numbers. *Journal of the Royal Society, Series A, 134*, 486–501.

Jennrich, R. I., & Robinson, S. M. (1969). A Newton-Raphson algorithm for maximum-likelihood factor analysis. *Psychometrika, 34*, 111–123.

Johnson, R., & Wichern, D. (1988). *Applied multivariate statistical analysis*, Englewood Cliffs, NJ: Prentice-Hall.

Kaiser, M. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika, 23*, 187–200.

Mullen, F. (1939). Factors in the growth of girls seven to seventeen years of age. Unpublished doctoral dissertation, Department of Education, University of Chicago.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*, 461–464.