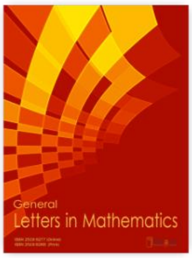

General Letters in Mathematics, Vol. 4, No.2 , April 2018, pp. 76 - 85
e-ISSN 2519-9277, p-ISSN 2519-9269
Available online at http:\\ www.refaad.com
https://doi.org/10.31559/glm2016.4.2.4

Principal Components and Factor Analysis of Fertility Differentials in Tanzania

Christopher Mbotwa[1] and Rocky R. Akarro[2]

[1] University of Dar es Salaam Mbeya College of Health and Allied Sciences, Mbeya, Tanzania
cmbotwa@gmail.com
[2] University of Dar es Salaam College of Social Sciences, Dar es Salaam, Tanzania
akarror@gmail.com

Abstract The objective of this paper is to apply principal components and factor analysis techniques in assessing factors associated with fertility differentials in Tanzania. The study utilized secondary data from 2010 Tanzania Demographic and Health Survey (2010 TDHS) dataset. Three factors were identified as the main factors associated to fertility differentials in Tanzania. The first factor was woman' education and awareness, the second factor was woman' demographic characteristics and the third factor was woman' economic status. Among those factors, woman' education and awareness was found to contribute more than all other factors in explaining fertility differentials in Tanzania. The study concluded that, in order to attain a desirable fertility level in the country, woman' education especially on the issue of family planning needs to be improved.

Keywords: *Eignvalues, Fertility, Loadings matrix, Principal Components, Tanzania.*

1 Introduction

In demography, fertility may be defined as the production of offspring as opposed to a physical capability to produce which is referred as fecundity. This is an important variable in population dynamics because the level and rate of fertility impact on the population growth and age pyramid structure.

High fertility is defined as five or more births per woman over the reproductive career [1]. High fertility countries are those countries whose Total Fertility Rate (TFR) is more than 3.2 children per woman [2]. According to [2] "high fertility countries are increasing in numbers and are concentrated in Africa (45 out of 66 high fertility countries) while those countries with low fertility are becoming more sparse geographically, and are mostly in Europe (only 31 out of 70 low fertility countries are from regions outside Europe)".

The high rate of population growth that results from high fertility levels makes provision of social services like education and health services very challenging, as these services need to expand with the growing population. When fertility decreases rapidly, children have higher access to education, people have greater assets accumulation, and there is greater use of preventive health services [3].

High fertility, which results in high rate of population growth, has a negative impact on the environment and natural resources [2,4]. The high rate of population growth has brought effects in deforestation and desertification, which mainly affect Sub-Saharan Africa due to dependency in agriculture [4].

According to [5], "postponement of first marriage, and divorce or widowhood is associated with the low rates of fertility". Gross national income, percentage of currently married women using modern contraceptive methods and female literacy rate has a significant impact in reducing fertility levels [6]. In Ghana and Nigeria, woman's education, woman's age at first marriage, woman's marital status, place of residence, wealth index of the household and use of contraceptive measures are the main determinants of high fertility level [7].

Though studies on causal factors of fertility are known, they do not identify the variables which account for large variation of fertility levels. Furthermore, these studies do not group variables according to their significance in explaining fertility levels and differentials; hence, it becomes difficult for policymakers to identify variables which contribute much to levels and variation of fertility in Tanzania.

This paper employs principal components and factor analysis techniques in assessing factors associated with fertility differentials in Tanzania using the 2010 TDHS data set. One advantage of this technique is that it seeks for the lower dimension, while preserving the large number of information from the original variables [8]. The paper identifies the variables which are responsible in explaining large variation of fertility in Tanzania by the use of factor loadings.

2 Materials and Methods

2.1 Materials

The paper utilized the Tanzania Demographic and Health Survey 2010 data set (2010 TDHS). The 2010 TDHS is the national representative of 10,300 households selected from 475 sample points. The 2010 TDHS data set contains information on fertility levels and preferences, marriage, sexual activities, awareness and use of family planning methods, maternal and child health, breastfeeding practices, nutritional and anaemia status of women and young children, childhood mortality, use of bed nets and anti-malaria, awareness and behaviour regarding Human Immunodeficiency Virus (HIV)/Acquired Immune Deficiency Syndrome (AIDS) and other STIs, female genital cutting, and adult and maternal mortality.

2.2 Data Analysis

Data analysis was done by STATA version 12. Principal components and factor analysis techniques were applied to extract factors associated to fertility differentials in Tanzania. Prior to Principal component analysis and factor extraction, correlation test and Kaiser Meyer-Olkin (KMO) and Bartlett's Test of Sphericity (BTS) were employed to test the suitability of the variables for principal components and factor analysis.

2.2.1 Principal Components Analysis

Consider the random vector $X' = (X_1, X_2, \dots, X_P)$ having the covariance matrix $\sum$ with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_P \geq 0$, then the principal components can be written as the linear combinations of these random variables such that they are uncorrelated and have maximum variance [9,10]. Consider the following linear combinations:

$$Y_1 = \alpha'_1 X = \alpha_{11} X_1 + \alpha_{12} X_2 + \cdots + \alpha_{1P} X_p$$
$$Y_2 = \alpha'_2 X = \alpha_{21} X_1 + \alpha_{22} X_2 + \cdots + \alpha_{2p} X_p$$
$$\vdots \qquad \qquad \vdots$$
$$Y_p = \alpha'_p X = \alpha_{p1} X_1 + \alpha_{p2} X_2 + \cdots + \alpha_{pp} X_p$$

With $\text{Var}(Y_i) = \alpha'_i \sum \alpha_i$ and $\text{Cov}(Y_i, Y_j) = \alpha'_i \sum \alpha_j$

Where by the values α_{ij}'s for $i, j = 1,2, \dots, p$ are the weights and the element of the eigenvector.

The principal components are those uncorrelated linear combinations $Y_1, Y_2, .., Y_P$ whose variances are as large as possible. The first principal component is the linear combination, which maximizes $Var(Y_1)$, but uncorrelated with the remaining components.
Since $Var(Y_1)$ can be maximized by multiplying any α_1 by some constants then, it is desirable to restrict attention to coefficient vectors of unit length. We thus, define the first principal component to be the linear combination of $Y_1 = \alpha'_1X$ that maximizes $Var(Y_1)$ subject to $\alpha'_1\alpha_1 = 1$. The second principal component is the linear combination of $Y_2 = \alpha'_2X$ that maximizes $Var(Y_2)$ subject to $\alpha'_2\alpha_2 = 1$ and $Cov(Y_1, Y_2) = 0$. The i^{th} Principal component is the linear combination of $Y_i = \alpha'_iX$ that maximizes $Var(Y_i)$ subject to $\alpha'_i\alpha_i = 1$ and $Cov(Y_i, Y_j) = 0$ for $j \neq i$.
If $\sum$ has the eigenvalue-eigenvector pairs $(\lambda_1, e_1), (\lambda_2, e_2), \ldots, (\lambda_p, e_p)$ where $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_p \geq 0$, then the i^{th} principal component can be written as;
$Y_i = \alpha_i'X = e_{i1}X_1 + e_{i2}X_2 + \cdots + e_{ip}X_p$
With this choice, $Var(Y_i) = \lambda_i$ for $i = 1,2, \cdots, p$ and $Cov(Y_i, Y_j) = 0$ for $i \neq j$.

2.2.2 Testing Suitability of the Variables for Principal Components Analysis

The first thing to do prior to principal components analysis is to test for the correlations and sampling adequacy. In order to carry out principal component analysis, each variable needs to be at least moderate correlated to one or more of the other variables. The correlation matrix was used to test for the correlations of variables. Variables that were not correlated with any of the other variables were dropped out from the analysis. Sampling adequacy was tested using KMO; and finally, Bartlett test was employed to test the null hypothesis that the correlation matrix was an identity matrix.

2.2.3 Extraction of Factors

Factors were extracted basing on the criteria of eigenvalues. Those factors with eigenvalues equal or larger than one were retained; the rest factors were dropped out. Factor Loadings matrix was constructed from retained factors and they were used to group variables responsible in explaining fertility differentials in Tanzania. The factors were sorted according to their contribution in explaining fertility differentials in Tanzania.

2.2.4 Definitions of Variables

The variables used in this paper are listed and defined in Table 1. The choice and categorization were mainly based on past studies such as [1,2].

Table 1: Definitions of variables used in this paper

Variable Name	Definition
1. Residence	Woman's type of place of residence (Rural or Urban)
2. Literacy	Ability to read and write of a woman aged 15-49 (literate or illiterate)
3. Wealth Index	Status of wealth of the household (poorest, poor, middle, richer, and richest)
4. Age-marriage	Woman's age at first marriage
5. Age-birth	Woman's age at first birth
6. Women-education	Education level for woman aged 15-49 (No education, primary incomplete, primary complete, and secondary +)

Table 1 (Continue)

Variable Name	Definition
7. Partner-education	Man's education level where a man is referred to one in marriage relationship with an interviewed woman (No education, primary incomplete, primary complete, and secondary +)
8. Women-occupation	Economic activity for which a woman is engaged (Not working, professional and services, business, agriculture, and others)
9. Partner-occupation	Economic activity for which a man is engaged where a man is referred to one in marriage relationship with an interviewed woman (Not working, professional and services, business, agriculture, and others)
10. Contraceptive use	Status and types of contraceptive methods used by woman aged 15-49 (No method, Traditional method, modern method)

3 Results and Discussion

3.1 Checking for Unique Variables

He we check if there are enough correlations to carry out principal component analysis and factor analysis. It is recommended to accept variables which are moderately correlated to one or more of the other variables ($r \geq 0.3$) [11,12,13]. The results show that out of ten variables, seven variables were at least moderately correlated with one of the other variables. Three variables, that is woman's occupation, partner's occupation and contraceptive use by method, were weakly correlated with all other variables so they were discarded prior to performing principal component analysis (Table 2).

Table 2: Pearson correlations matrix

Variables	1	2	3	4	5	6	7	8	9	10
1. Residence	1.00									
2. Literacy	-0.19	1.00								
3. Wealth-index	*-0.52*	*0.32*	1.00							
4. Age-birth	-0.06	0.06	0.06	1.00						
5. Age-marriage	-0.09	0.10	0.08	*0.66*	1.00					
6. women-education	-0.22	*0.78*	*0.36*	0.11	0.14	1.00				
7. Partner-education	-0.24	*0.34*	*0.37*	0.06	0.09	*0.40*	1.00			
8. women-occupation	0.00	0.05	0.01	-0.08	-0.04	0.01	0.00	1.00		
9. Contraceptive use	-0.10	0.13	0.12	-0.03	0.04	0.13	0.15	0.07	1.00	
10. Partner-occupation	0.03	-0.02	-0.09	-0.02	-0.01	-0.04	-0.13	0.06	0.03	1.00

3.2 KMO and BTS

KMO and BTS are used in assessing the strength of the relationship. The BTS compares the correlation matrix with identity matrix (a matrix of zero correlations). [14] noted that "BTS can be used to test the null hypothesis that the sample was randomly drawn from a population in which the correlation matrix was an identity matrix". KMO is used as the measure of sampling adequacy. [15] recommended the acceptable KMO value in order to continue with the analysis is the value greater than 0.5. For BTS, the associated p-value should be less than the significance level (0.05). In this case, KMO Value is 0.631, and P-value of BTS is 0.000 (Table3).

Table 3: KMO and BTS

KMO Measure of Sampling Adequacy.		.631
BTS	Approx. Chi-Square	54,240.534
	Degree of freedom	21
	P-value	0.000

3.3 Extraction of Principal Components

From p variables, p components can be extracted. This involves finding the solution to p equations with p unknowns. According to [9], "The variance in the correlation matrix is repackaged into p eigenvalues and each eigenvalue represents the amount of variance that has been captured by one component".

Each of the p variable's variance is standardized to one. Each factor's eigenvalue may be compared to 1 to see how much more (or less) variance it represents than does a single variable. According to [10] "The principal components extraction will produce p components which in the aggregate account for all of the variance in the p variables". Table 4 shows the eigenvalues and the proportion of variance explained by the eigenvalue. Table 5 gives the eigenvectors formed from the eigenvalues (coefficients of principal components). Three principal components were found to have eigenvalues greater than one, and the cumulative percentage of variance explained by these three principal components is 75.4%.

Table 4: Eigenvalues and percentage of variance

Component	Eigenvalues		
	Total	Percentage of Variance	Cumulative Percent
1	2.597	37.102	37.102
2	1.582	22.602	59.703
3	1.099	15.699	75.403
4	0.705	10.074	85.477
5	0.457	6.532	92.009
6	0.34	4.86	96.869
7	0.219	3.131	100

Table 5: Initial principal components (eigenvectors)

Variable	Comp1	Compp2	Comp3	Comp4	Comp5	Comp6	Comp7
Residence	-0.3485	0.1574	0.6422	0.27	0.6038	-0.0217	0.0583
Literacy	0.4432	-0.1116	0.4811	-0.343	-0.0923	0.5292	0.3917
Wealth_index	0.4318	-0.2063	-0.4153	-0.0658	0.7663	0.0019	0.0832
Age_marriage	0.2487	0.656	-0.0551	0.028	-0.0512	-0.4281	0.564
Age_birth	0.2283	0.6692	-0.0586	0.0187	0.0798	0.435	-0.5483
Women_education	0.477	-0.1204	0.4185	-0.1724	-0.0374	-0.5798	-0.4643
Partner_education	0.3926	-0.1659	0.0444	0.8799	-0.1713	0.1039	0.0438

3.4 Rotation of the Components

In order to obtain meaningful and interpretable principal components, the rotation was done to the Initial principal components. The method of rotation used was varimax with Kaiser normalization method [9]. The principal components after rotation are given in Table 6.

Table 6: Principal components after rotation

Variable	Comp1	Comp2	Comp3	Unexplained
Residence	0.1148	-0.0184	*-0.7384*	0.2167
Literacy	*0.6583*	0.0032	-0.0832	0.2303
Wealth index	0.1013	-0.0181	*0.6253*	0.2707
Age at first marriage	0.0138	*0.7035*	0.0101	0.1684
Age at first birth	-0.0063	*0.7095*	-0.0026	0.166
Woman's education	*0.6456*	0.0111	-0.0131	0.2053
Partner's education	*0.3551*	-0.0297	0.238	0.5557

Thus, ignoring the variables with low coefficients (coefficients whose absolute values are less than 0.3) to each component we obtain the variables that represent each of the components in the linear combinations.

Comp 1 = 0.6583(Literacy) + 0.6456(Woman's education) + 0.3551(Partner's education)

Comp 2 = 0.7035(Age at first marriage) + 0.7095(Age at first birth)
Comp 3 = -0.7384(Type of place of residence) + 0.6253(Wealth index)

3.5 Meaning of the Coefficients of Principal Components

The coefficients of the principal components are very useful in identifying the relative importance of the original variables to the principal components. Coefficients of the principal components help a researcher to know which variables are important to a particular component. The variables whose coefficients are greater than or equal to 0.5 are considered as important to that principal component.

In this case, literacy and woman's education are important to the first principal component with coefficients 0.66 and 0.65 respectively. Thus, the first principal component is associated with the woman's education and awareness**.** Age at first marriage and age at first birth are important to the second component with coefficients 0.70 and 0.71 respectively. Accordingly, the second principal component is associated with the woman's demographic characteristics**.** Type of place of residence and Wealth index are important to the third principal component with coefficients -0.74 and 0.63 respectively. Thus, the third principal component is associated with woman's economic status**.**

It can be noted that partner's education is not as much as important to any of the three components. Literacy and woman's education are almost equally important to the first principal component; likewise, age at first marriage and age at first birth to the second component. In the third component, Type of place of residence is more important than wealth index because the former variable has a larger coefficient as compared to the later variable.
The results in coefficients of the principal components give similar interpretation as those obtained in the factor loadings. Moreover, naming of the components is accomplished by the rotated factor loadings although those names have been used in the interpretation of the coefficients of principal components.

3.6 Results on Factor Analysis

Three factors (woman's education and awareness, woman's demographic characteristics and woman's economic status) were extracted using the principal components analysis techniques. The factors formed were rotated using varimax rotation approach and then named according to their underlying groups.

3.7 Factor Loadings Matrix

Loading matrix is one of the important components in FA as it shows the correlation of the factors to the original variables. The entries in this matrix, loadings, are correlations between the factors and the variables. Factor loadings are used to assess on how well the original variables are explained by the underlying factors attained. According to [16], "The loadings that are more than 0.5 are typically considered strong, between 0.3 and 0.5 are moderate, and less than 0.3 are typically considered as weak".

3.8 Loadings Matrix for Unrotated Factors

The results show that factor 1 is strongly correlated with five variables (woman's education, literacy, wealth index, partner's education, and type of place of residence); moderately correlated with age at first marriage, and weakly correlated with age at first birth.
Factor 2 is strongly correlated with only two variables (age at first birth and age at first marriage), and it is weakly correlated with all other remaining variables. The third factor is strongly correlated to only one variable (type of place of residence); moderately correlated with three variables (woman's education, literacy, and wealth index); and it is weakly correlated with the rest variables.

Table 7 shows the results of unrotated factors (the italicized numbers show the variables which are at least moderately correlated with the respective factors).
However, naming of unrotated factor loadings is difficult because almost all variables are at least moderately correlated to more than one factor. Moreover, it is difficult to obtain meaningful names for the factors attained using unrotated factor loadings. Therefore, in order to obtain meaningful names and have common factor for each variable, rotation of the factor loadings is compulsory. The results on the rotated factor loadings are discussed in the succeeding section.

Table 7: Components matrix for unrotated coefficients

Variables	Factor		
	1	2	3
Woman's education	*0.806*	-0.136	*0.436*
Literacy	*0.759*	-0.183	*0.485*
Wealth index	*0.683*	-0.184	*-0.478*
Partner's education	*0.631*	-0.153	-0.025
Age at first birth	0.295	*0.862*	-0.021
Age at first marriage	*0.337*	*0.845*	-0.009
Type of place of residence	*-0.552*	0.123	*0.667*

3.9 Loadings Matrix for Rotated Factors

In order to have a meaningful interpretation, we need to look at the rotated loadings, where each variable is strongly associated to only one of the factor, and weakly associated with the other remaining factors. Table 8 shows the results of the rotated factor loadings (italicized numbers show the variables which are significant to the respective factors).

In this case, factor 1 is strongly correlated with literacy and woman's education. In additional, this factor is moderately correlated with partner's education. Thus, factor 1 can be named as woman's education and awareness, and this is represented well by woman's literacy and education attainment. Factor 2 is strongly correlated with age at first marriage and age at first birth. Accordingly, it can be named as woman's demographic characteristics. Factor 3 is strongly correlated with type of place of residence and wealth index, consequently this factor form a group woman's economic status.

Table 8: Components matrix for rotated factors

Variables	Factor		
	1	2	3
Literacy	*0.914*	0.027	0.092
Woman's education	*0.91*	0.087	0.149
Partner's education	*0.499*	0.03	0.416
Age at first birth	0.031	*0.91*	0.028
Age at first marriage	0.074	*0.906*	0.047
Type of place of residence	-0.04	-0.049	*-0.872*
Wealth index	0.269	0.024	*0.81*

3.10 Communalities

According to [9], "Communality is the amount of the variable's variance that is accounted by the factor, and because the loadings are correlations between variables and factors, and the factors are orthogonal; then, a variable's communality represents the R^2 of the variable predicted from the factors".

The communalities for the variables that load well in the first factor (woman's education and awareness) are woman's education, 0.859; literacy, 0.845; and partner's education, 0.422. This shows that most of the variance in the factor 1 is accounted by woman's education and literacy while partner's education explains a moderate variation. These results lead to the same meaning as those in the factor loadings; partner's education shows a weak correlation with the first factor.

The communalities of woman's demographic characteristics variables (age at first marriage and age at first birth) are 0.828 and 0.8311 respectively. These are the variances accounted by these variables in the three factors formed. Woman's economic status variables account for less variation as compared to others (Type of place residence, 0.764; and Wealth index, 0.729).
Therefore, woman's education and awareness variables are the most important variables that explain the large variation of fertility in Tanzania. Other variables are woman's demographic characteristics variables, and woman's economic status variables. Strictly speaking, woman's education, literacy, age at first marriage, age at first birth, wealth index, and type of place of residence are the most important variables in explaining fertility differentials in Tanzania. Table 9 shows the results of the communalities.

Table 9: Communalities of the variables used in factor analysis

	Initial	Extraction
Type of place of residence	1	*0.764*
Literacy	1	*0.845*
Wealth index	1	*0.729*
Age at first birth	1	*0.831*
Age at first marriage	1	*0.828*
Woman's education	1	*0.859*
Partner's education	1	0.422

3.11 Factor Scores

Factor scores are defined as estimates of underlying factor values for each observation. The results of the factor scores have been displayed in matrix form as shown in Table 10.

3.12 Factor Scores Loadings Matrix

This matrix is used to show the loadings of factor scores on variables. It indicates how the standardized factors load on the original variables. In this case, it was found that literacy and woman's education have high loading in the first factor score; age at first marriage and age at first birth load well on the second factor; and type of place of residence and wealth index load well on the third factor. Partner's education does not show high loading in any of the three factor scores (Table 10). This indicates that partner's education is not as much as important to any of these three factors. The factor scores attained may save as inputs for further applications such as regression and clustering analysis.

Table 10: Factor score coefficients matrix

Variable	Factor		
	1	2	3
Residence	0.192	0.004	*-0.619*
Literacy	*0.517*	-0.038	-0.158
Wealth index	-0.04	-0.03	*0.519*
Age-birth	-0.04	*0.555*	-0.02
Age-marriage	-0.019	*0.549*	-0.017
Women-education	*0.497*	-0.003	-0.118
Partner-education	0.192	-0.024	0.178

4 Conclusion and Recommendations

4.1 Conclusion

Three components formed from seven variables (woman's literacy, woman's education, partner's education, age at first marriage, age at first birth, type of place of residence, and wealth index) were sufficient in explaining the total variance of socio-economic and demographic variables associated with fertility differentials in Tanzania.

Each of the three components formed was strongly correlated with two of the variables. The first principal component was strongly correlated with woman's literacy and education. The second principal component was strongly correlated with age at first marriage and age at first birth. The third principal component was strongly correlated with the type of place of residence and wealth index.

Three underlying factors were attained by the use of principal components method. The first factor which corresponds to the first component was named as woman's education and awareness. The second factor which corresponds to the second component was named as woman's demographic characteristics. The third factor was named as woman's economic status. All original variables except partner's education, had shown high loading to one of the three factors formed. This shows that three factors formed were sufficient in explaining factors associated with fertility differentials in Tanzania.

Fertility differentials seem to be determined mostly by woman's education and awareness because it account for large variance compared to other factors (woman's demographic characteristics and woman's economic status). This may be because other factors are predominately determined by woman's education and awareness.

4.2 Recommendations

This paper recommends that Tanzania Government should find ways of improving woman's education in moral and ethical aspects. This is because woman's education and awareness is the component that explains most of the variance among all other factors associated with fertility differentials in Tanzania. Further researches should concentrate on spatial analysis of fertility levels and differentials in Tanzania to identify areas in the country at high risk of high fertility rate.

Acknowledgements

The authors would like to thank Demographic and Health Survey (DHS) program for giving us the access to their dataset and a reviewer for the helpful comments and suggestions.

References

[1] National Bureau of Statistics (NBS) [Tanzania] and ICF Macro, Tanzania Demographic and Health Survey 2010. Dar es Salaam, Tanzania: ICF Macro, (2011).

[2] United Nations Department of Economic and Social Affairs, Population Division: World Fertility Report 2013: Fertility at Extremes, United Nations Publications, (2010).

[3] J. Bongaarts, J. Cleland, J.W. Townsend, J.T. Bertrand & M.D. Gupta, Family planning programs for the 21st century, New York: Population Council, (2012).

[4] World Health Organization. World health statistics 2010. World Health Organization, (2010).

[5] A. Mturi & A. Hinde, Fertility levels and differentials in Tanzania. In Proceedings of the United Nations Workshop on Prospects for Fertility Decline in High Fertility Countries, Department of Economics and Social Affairs, New York, (2001).

[6] A.J.M.A Sufian, Multivariate Analysis of Cumulative Fertility Using Secondary Data, International Journal of Business and Social Science, 4(8) (2013).

[7] O. Olatoregun, A.F. Fagbamigbe, O.J. Akinyemi, O.B. Yusuf & E.A. Bamgboye, A comparative analysis of fertility differentials in Ghana and Nigeria: original research article, African journal of reproductive health, 18 (3) (2014), 36-47.

[8] H. L. Shang , H. Booth & R. Hyndman, Point and interval forecasts of mortality rates and life expectancy: A comparison of ten principal component methods, Demographic Research, 25 (2011), 173-214, https://doi.org/10.4054/demres.2011.25.5

[9] R.A. Johnson & D.W. Wichern, Applied Multivariate Statistical Analysis, Prentice Hall: Englewood Cliffs, NJ, (2007).

[10] A.C. Rencher, Methods of multivariate analysis. New York: John Wiley & Sons, (2003).

[11] J.C. Anderson & D.W. Gerbing, The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis, Psychometrika, 49(2)(1984), 155-73, https://doi.org/10.1007/bf02294170

[12] J.O. Kim & C.W. Mueller, Factor analysis: Statistical methods and practical issues, Sage, (1978).

[13] C.D. Dziuban & E.C. Shirkey, When is a correlation matrix appropriate for factor analysis? Some decision rules, Psychological bulletien, 81(6)(1974), 358-361, https://doi.org/10.1037/h0036316

[14] K.F. Widaman, Common factor analysis versus principal component analysis: Differential bias in representing model parameters?, Multivariate Behavioral Research, 28(3)(1993), 263-311, https://doi.org/10.1207/s15327906mbr2803_1

[15] H.F. Kaiser, An index of factorial simplicity, Psychometrika, 39(1)(1974),31-36, https://doi.org/10.1007/bf02291575

[16] T.H. reiblmaier & P. Filzmoser, Exploratory factor analysis revisited: How robust methods support the detection of hidden multivariate data structures in IS research, Information & management, 47(4)(2010), 197-207, https://doi.org/10.1016/j.im.2010.02.002